AD-A114 536   WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER   F/G 12/1
SIMULTANEOUS SIMILARITY OF MATRICES.(U)
MAR 82 S FRIEDLAND   DAAG29-80-C-0041
UNCLASSIFIED   MRC-TSR-2345   NL

1 OF 2
AD
A114536

2

AD A114536

MRC Technical Summary Report #2345

SIMULTANEOUS SIMILARITY OF MATRICES

Shmuel Friedland

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

March 1982

Received October 2, 1981

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

SIMULTANEOUS SIMILARITY OF MATRICES

Shmuel Friedland[†]

Technical Summary Report #2345
March 1982

ABSTRACT

In this paper we solve completely and explicitly the long standing problem of classifying pairs of $n\times n$ complex matrices (A,B) under ~~the~~ a simultaneous similarity, ~~(TAT^{-1},TBT^{-1})~~. Roughly speaking, the classification decomposes to a finite number of steps. In each step we consider an open algebraic set, $\mathcal{M}^0_{n,2,r,\rho} \subset M_n \times M_n$ (M_n - the set of $n\times n$ complex valued matrices). Here r and ρ are two positive ~~integers.~~ Then we construct a finite number of rational functions ~~$\phi_1,\cdots,\phi_s$~~ in the entries of A and B whose values are constant on all pairs similar ~~in $\mathcal{M}_{n,2,r,\rho}$~~ to (A,B). The values of the functions $\phi_i(A,B)$, $i = 1,\cdots,s$, determine a finite number (~~at most~~ $\kappa(n,2,r)$) of similarity classes, in $\mathcal{M}_{n,2,r,\rho}$. Let S_n be the subspace of complex symmetric matrices in M_n. For $(A,B) \in S_n \times S_n$ we consider the similarity class (TAT^t, TBT^t) where T ranges over all complex orthogonal matrices. Then the characteristic polynomial $|\lambda I - (A+xB)|$ determines a finite number of similarity classes for almost all pairs $(A,B) \in S_n \times S_n$.

AMS(MOS) Subject Classification - 15A21, 15A90

Key Words: Simultaneous similarity, invariant functions, symmetric matrices

Work Unit Number 1 - Applied Analysis

[†] Institute of Mathematics, Hebrew University, Jerusalem; Mathematics Research Center, University of Wisconsin-Madison and Department of Mathematics, University of Wisconsin-Milwaukee.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

Consider the following linear singular differential system

$$(1) \qquad \frac{dx}{dt} = (A + B/t)x.$$

Here $x = (x_1,\cdots,x_n)^t$ is a column vector and $A = (a_{ij})_1^n$ and $B = (b_{ij})_1^n$ are $n\times n$ complex valued matrices. A linear change $y = Tx$, transforms (1) to

$$(2) \qquad \frac{dy}{dt} = (A_1 + B_1/t)y, \quad A_1 = TAT^{-1}, \quad B_1 = TBT^{-1}.$$

Thus we are led to study the properties of pairs of matrices (A_1,B_1) which are simultaneously similar to (A,B). It can be shown that more complicated singular differential systems under more general transformations will lead again to a certain similarity class of pairs of matrices.

We now give another example of simultaneous similarity. Let A and B be real symmetric. Then A and B can represent a potential and kinetic energy of some physical system. By changing to another orthonormal basis we shall get another pair (A_1,B_1)

$$(3) \qquad A_1 = TAT^t, \quad B_1 = TBT^t$$

for the corresponding orthogonal matrix T. The class of all such symmetric pairs (A_1,B_1) gives all representations of the same physical system. Thus to classify such similarity class is equivalent to classifying the physical system up to a choice of an orthonormal basis.

In this paper we solve completely the first mentioned problem. The classification of the similarity classes (TAT^{-1},TBT^{-1}) decomposes to a finite number of steps. In each step we consider an algebraic variety V (a set of points defined by polynomial equations) of pairs of matrices. We construct a set of invariant functions, sort of "generalized eigenvalues" of the pairs (A,B) lying in V. Those "eigenvalues" classify all similarity classes of (A,B) except for some algebraic subvariety W in V. We now repeat the process for W.

In the case of symmetric pairs (A,B), we do as follows. With each such pair we associate the characteristic polynomial $p(\lambda,x) = |\lambda I - (A+xB)|$. Clearly, any similar pair (A_1,B_1) has the same characteristic polynomial. In this paper we show that for almost all symmetric pairs (A,B) the characteristic polynomial $p(\lambda,x)$ determines the orthogonal similarity class up to a finite number of possibilities. We strongly believe that our results would have some applications to the invariants of singular differential systems as well as to other areas, in particular mathematical physics.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

TABLE OF CONTENTS

SIMULTANEOUS SIMILARITY OF MATRICES

Shmuel Friedland[†]

0. Introduction.

Let M_n denote the s : of $n \times n$ complex valued matrices and G_n the group of all invertible matrices in M_n. It is a classical problem to classify the similarity classes (orbits) of m+1-tuples of matrices $(A_0, \cdots, A_m)$ under the action of G_n,

$$\mathrm{orb}(A_0, \cdots, A_m) = \{(B_0, \cdots, B_m), \quad B_i = TA_iT^{-1}, \quad i = 0, \cdots, m, \ T \in G_n\}. \tag{0.1}$$

See Gelfand [1970], Gelfand-Ponomarev [1969], Brenner [1975], Nathanson [1980], Processi [1976] and Friedland [1980] for certain problems in which the classification of such orbits needed, for various results on this problem and additional references. It is known that the classification of similarity classes of m+1 tuples can be reduced to the classification of simultaneous similarity of pairs of matrices (A,B). In fact one can assume that A and B are commuting and even nilpotent. See Gelfand-Ponomarev [1969] and Nathanson [1980]. Therefore in certain cases, for the simplicity of the exposition we are dealing with the simultaneous similarity of pairs of matrices (m=1). In cases when the choice m=1 does not simplify the treatment of the problem we are dealing with an arbitrary m. In Friedland, [1981] we outlined a general procedure for classifying the orbits of a given algebraic group acting on an irreducible variety over any algebraically closed field. For the problem of simultaneous similarity our procedure works as follows. In the step number i we are given an irreducible variety V in

$$\mathcal{M}_{n,m+1} = \underbrace{M_n \times \cdots \times M_n}_{m+1} \tag{0.2}$$

† Institute of Mathematics, Hebrew University, Jerusalem; Mathematics Research Center, University of Wisconsin-Madison and Department of Mathematics, University of Wisconsin-Milwaukee.

Sponsored by the United States Army Under Contract No. DAAG29-80-C-0041.

which is invariant under the action of G_n. That is if V contains a tuple $(A_0,\cdots,A_m)$ then V contains $\mathrm{orb}(A_0,\cdots,A_m)$. Let $[V]$ and (V) be the ring of polynomials and its field of quotients whose variables are $(m+1)n^2$ entries of $A_k = (a_{ij}^{(k)})$, $k = 0,\cdots,m$ and whose values are restricted to V. A rational function $f \in (V)$ is called invariant if f is constant on all orbits lying in V. Denote by $[V]^G$ and $(V)^G$ the subring and subfield of invariant functions in $[V]$ and (V) respectively. Since $(V)^G$ is a subfield of rational functions in $(m+1)n^2$ variables $(a_{11}^{(0)},\cdots,a_{nn}^{(0)},\cdots,a_{11}^{(m)},\cdots,a_{nn}^{(m)})$ it is known that $(V)^G$ is finitely generated. That is there exist q invariant functions $\chi_1,\cdots,\chi_q \in (V)^G$ such that any $\chi \in (V)^G$ can be expressed as a rational function in $\chi_1,\cdots,\chi_q$. See for example Fogarty [1969, p. 69]. Then there exists an invariant (strict) subvariety $W \subseteq V$ such that the functions $\chi_1,\cdots,\chi_q$ are defined on each point of the open set $V-W$ and their values determine the orbit of $(A_0,\cdots,A_m)$. The orbits in $V-W$ are characterized by two numbers their dimensions and their degrees as irreducible algebraic varieties in $\mathcal{M}_{m,m+1}$. So in the next stage one has to classify orbits on W and the process ends in a finite number of steps. In Section 1 we identify the sets $V-W$ with open algebraic sets $\mathcal{M}^0_{n,m+1,r,\rho}$ which are characterized by two integers r and ρ. In fact n^2-r is the dimension of any orbit in $\mathcal{M}_{n,m+1}$ and ρ is another integer which has a relatively simple characterization. Roughly speaking ρ is the number of linearly independent polynomial equations which determine the orbit of $(A_0,\cdots,A_n)$. We show explicitly how to find a set of invariant functions $\phi_1,\cdots,\phi_s$ in $(\mathcal{M}_{n,m+1,r,\rho})^G$ such that any orbit in $\mathcal{M}^0_{n,m+1,r,\rho}$ is determined by the value of these functions up to a finite number $\kappa(n,m,r)$ of them.

To understand the complexity of the general problem we give a complete classification of pairs of 2×2 matrices (A,B) under the simultaneous similarity (TAT^{-1},TBT^{-1}). We do it directly without using the results of

Section 1. Our classification then is reduced to three steps. In the first step the coefficients of the characteristic polynomial $|\lambda I - (A+xB)|$ determine the orbit of (A,B) completely as long as $|\lambda I - (A+xB)|$ is irreducible over the ring $\mathbb{C}[\lambda,x]$. Let U be the set of all pairs (A,B) such that $|\lambda I - (A+xB)|$ splits to a product of two linear factors. Then U is given by a single equation

$$U = \{(A,B), \{2tr(A^2)-[tr(A)]^2\}\{2tr(B^2)-[tr(B)]^2\} = [2tr(AB) - tr(A)tr(B)]^2.$$

It follows that U is an irreducible variety of codimension 1. Let V be the set of all pairs (A,B) such that $|\lambda I - (A+xB)|$ is a square of a linear factor. That is, for any x $A+xB$ has a double eigenvalue. Then V is given by a set of these equations

$$V = \{(A,B),\ 2tr(A^2) = [tr(A)]^2,\ 2tr(B^2) = [tr(B)]^2,\ 2tr(AB) = tr(A)tr(B)\}.$$

So V is an irreducible variety of codimension 3. We show that on $U-V$ the characteristic polynomial of $|\lambda I - (A+xB)|$ determine 3 distinct orbits. In order to distinguish between these 3 orbits we have to introduce rational invariant functions in U

$$\alpha(A,B) = \frac{(b_{11}-b_{22})a_{12}a_{21} + a_{22}a_{11}b_{21} - a_{11}a_{21}b_{12}}{a_{12}b_{21} - a_{21}b_{12}}$$

$$\beta(A,B) = \frac{(a_{11}-a_{22})b_{12}b_{21} + b_{22}b_{12}a_{21} - b_{11}b_{21}a_{12}}{b_{12}a_{21} - b_{21}a_{12}}.$$

Those functions are defined on non-commuting pairs (A,B) in U. The orbit of the commuting pair $(A,B) \in V$ is determined by its characteristic polynomial. It is easy to see that $[V]^G$ is generated by $tr(A)$ and $tr(B)$. So the transcendence degree of the quotient field of $[V]^G$ is 2. However, the transcendence degree of $(V)^G$ is 3 since we have an additional rational invariant function on $\gamma(A,B) = a_{12}/b_{12} = a_{21}/b_{21}$.

Then $\mathrm{tr}(A)$, $\mathrm{tr}(B)$ and $\gamma(A,B)$ (or $1|\gamma(A,B)$) determine the orbit of (A,B) as long as A and B are not the scalar matrices. In the last case $\mathrm{tr}(A)$ and $\mathrm{tr}(B)$ determine $\mathrm{orb}(A,B)$. This shows that in the classification process carried out in Section 1 we must at a certain stage consider rational invariant functions in V. We now list briefly the contents of the rest of the paper. Sections 3-5 are technical and needed in the sequel. They deal with polynomial maps $p:\mathbf{C}^{\mu} \to \mathbf{C}^{\nu}$ and some special properties of certain algebraic functions in two variables. In Section 6 we deal with irreducible paris (pencils) (A,B). These are the pairs for which $|\lambda I - (A+xB)|$ is an irreducible polynomial over $\mathbf{C}[\lambda,x]$. The orbits of these pairs are determined completely by the generators of the ring of the invariant polynomials on $M_n \times M_n$. According to Processi [1976] these generators can be picked to be of the form $\mathrm{tr}(A^{i_1}B^{j_1},\cdots,A^{j_m}B^{j_m})$, and we described a procedure to determine the orbit. We showed that the transcendence degree of $[M_n \times M_n]^G$ is $n(n+1)$ however we were not able to find a transcendent basis in $[M_n \times M_n]^G$.

The rest of the paper is devoted to the study of $m+1$ tuple $(A_0,\cdots,A_m)$ of symmetric matrices under the action of complex orthogonal group O_n. In this context

$$\mathrm{sorb}(A_0,\cdots,A_m) = \{(B_0,\cdots,B_m),\ B_i = QA_iQ^t,\ i = 0,\cdots,m, Q \in O_n\}.$$

This case is of importance in mathematical physics at least for $m=1$. In that case we can interpret A_0 and A_1 as the potential and the kinetic energies. The orbit of (A_0,A_1) will correspond to the representations of the potential and the kinetic energies in different orthogonal bases. Usuallly A_0 and A_1 are real, however in some instances one considers the complex case too. See for example Moiseyev-Friedland [1980]. Let S_n be the set of $n\times n$ complex symmetric matrices and put

$$\mathcal{S}_{n,m+1} = \underbrace{S_n \times \cdots \times S_n}_{m+1}.$$

Denote by $[S_{n,m+1}]^0$ the set of invariant polynomials in the entries $(A_0,\cdots,A_m) \in S_{n,m+1}$ under the action of O_n. In Section 7 we show that the transcendent degree of $[S_{n,m+1}]^0$ is $n[(n+1)m+2]/2$ for $m=1$. The general case is discussed in Section 11. We also find a simple transcendence basis in this ring. Unfortunately, this basis is not symmetric in $A_0,\cdots,A_m$. Let $A(x) = \sum_{i=0}^{m} A_i x^i$ and consider $p(\lambda,x) = |\lambda I-A(x)|$ the characteristic polynomial of $A(x)$. Clearly, the coefficients of $p(\lambda,x)$ belong to $[S_{n,m+1}]^0$. The number of these coefficients is $n[(n+1)m+2]/2$! This suggests (as in 2×2 use) that these coefficients form a transcendence basis in $[S_{n,m+1}]^0$. This is indeed the case and Sections 8-11 are devoted to prove this result.

We also determine a class of characteristic polynomials $p(\lambda,x)$ which determine the orbit of $(A_0,\cdots,A_m)$ up to a finite number and we estimate from above the number of distinct orbits corresponding to $p(\lambda,x)$. In Section 8 we study polynomial matrices $A(x)$ with constant eigenvalues. In Section 9 we show that if $A,B \in S_n$, $A+xB$ has constant n distinct eigenvalues then $B = 0$ for $n \leq 4$. This result implies that for $n \leq 4$ the polynomial $p(\lambda,x) = |\lambda I - (A+xB)|$ determines at most

$$M(n) = \prod_{i=1}^{n} i!/2^{n-1}$$

distinct orbits of (A,B) (and generally this is the number of distinct orbits) on condition that $p(\lambda,x)$ is non-degenerate. That is $p(\lambda,x) = 0$ does not have a multiple root for all x.

For $n \geq 5$ this result does not apply. Section 10 deals with a general symmetric polynomial matrix $A(x)$. We show that if $p(\lambda,x) = |\lambda I - A(x)|$ is a non-degenerative polynomial and for each finite or infinite x each λ root of $p(\lambda,x) = 0$ is either single or double then there are at most $2^{(n-1)(mn-1)}$ distinct orbits corresponding to $m+1$ symmetric tuples

$(A_0, \cdots, A_m)$ such that

$$p(\lambda, x) = |\lambda I - \sum_{i=0}^{m} A_i x^i|.$$

The proof of this result is non-trivial and lengthy. The basic idea is to use the theory of one complex variable, in particular the Liouville theorem. The last section is devoted to various remarks, comments and conjectures.

1. Classification of similarity classes of tuples of matrices.

Let $A \in M_n$. As usual denote by $|A|$ the determinant of A. In what follows we adopt the notation of Marcus-Minc [1964].

Denote by $Q_{k,n}$ the set of strictly increasing sequences of integers $\alpha = (\alpha_1,\cdots,\alpha_k)$, $1 \le \alpha_1 < \cdots < \alpha_k \le n$. For any rectangular matrix A we denote by $A[\alpha|\beta]$ the submatrix generated by the rows $\alpha = (\alpha_1,\cdots,\alpha_k)$ and columns $\beta = (\beta_1,\cdots,\beta_\ell)$. In case that $k = \ell$ $|A[\alpha|\beta]|$ denote the appropriate minor of A. Also $A[i,j]$ denotes the (i,j) entry of A. For $A,B \in M_n$ let $L(B,A)$ be the following operator on M_n

$$L(B,A)X = BX-XA, \quad X \in M_n.$$

In tensor notation $L(B,A)$ is represented by the matrix $I \otimes B - A^t \otimes I$. Here by A^t we denote the transposed matrix of A. See for example Marcus-Minc [1964, p. 8]. Let $\nu(B_0,\cdots,B_m,A_0,\cdots,A_m)$ be the dimension of the subspace of matrices satisfying

$$B_iX - XA_i = 0, \quad i = 0,\cdots,m. \tag{1.1}$$

For $B_i = A_i$, $i = 0,\cdots,m$ denote the dimension of this subspace by $\nu(A_0,\cdots,A_m)$. Note that $\nu(A_0,\cdots,A_m) \ge 1$ since $X = I$ is a solution of (1.1) for $B_i = A_i$, $i = 0,\cdots,m$. Denote by $L(B_0,\cdots,B_m,A_0,\cdots,A_m)$ an $(m+1)n^2 \times n^2$ matrix composed of the submatrices $L(B_i,A_i)$, $i = 0,\cdots,m$. Clearly

$$\operatorname{rank}L(B_0,\cdots,B_m,A_0,\cdots,A_m) = n^2-\nu(B_0,\cdots,B_m,A_0,\cdots,A_m)\,. \tag{1.2}$$

Let $A,B \in M_n$. Suppose that $B = TAT^{-1}$. Then $L(A,X)$ and $L(X,A)$ are correspondingly similar to $L(B,X)$ and $L(X,B)$ for any $X \in M_n$

$$\begin{aligned} I \otimes B - X^t \otimes I &= (I \otimes T)(I \otimes A - X^t \otimes I)(I \otimes T^{-1}) \\ I \otimes X - B^t \otimes I &= ((T^t)^{-1} \otimes I)(I \otimes X - A^t \otimes I)(T^t \otimes I) \end{aligned} \tag{1.3}$$

So if $T(A_0,\cdots,A_m)T^{-1} = (B_0,\cdots,B_m)$ we have

$$L(B_0,\cdots,B_m,X_0,\cdots,X_m) =$$

$$\operatorname{diag}\{I\times T,\cdots,I\times T\}L(A_0,\cdots,A_m,X_0,\cdots,X_m)(I\times T^{-1}) \tag{1.4}$$

$$L(X_0,\cdots,X_m,B_0,\cdots,B_m) =$$

$$\operatorname{diag}\{(T^t)^{-1}\times I,\cdots,(T^t)^{-1}\times I\}L(X_0,\cdots,X_m,A_0,\cdots,A_m)(T^t\times I).$$

Here by $\operatorname{diag}\{A_0,\cdots,A_m\}$ we mean the block diagonal matrix with matrices $A_0,\cdots,A_m$ on the main diagonal. By choosing $X_i = A_i$ in the first equality in (1.4) we get

$$\nu(B_0,\cdots,B_m,A_0,\cdots,A_m) = \nu(A_0,\cdots,A_m) \tag{1.5}$$

if $(A_0,\cdots,A_m)$ and $(B_0,\cdots,B_m)$ are simultaneously similar. In general, the equality (1.5) does not imply the similarity of $(A_0,\cdots,A_m)$ and $(B_0,\cdots,B_m)$ even in the case $m=0$. See Friedland [1980]. However (1.5) implies the similarity of $(B_0,\cdots,B_m)$ and $(A_0,\cdots,A_m)$ provided that $(B_0,\cdots,B_m)$ lies in some open set U containing $(A_0,\cdots,A_m)$.

<u>Theorem</u> 1.6. Let $A_0,\cdots,A_m \in M_n$ and put $\nu = \nu(A_0,\cdots,A_m)$, $r = n^2-\nu$. Assume that $\alpha \in Q_{r,(m+1)n^2}$, $\beta \subset Q_{r,n^2}$ satisfy the following assumptions

(i) there exists $B_0,\cdots,B_m \in M_n$ such that

$$|L(B_0,\cdots,B_m,A_0,\cdots,A_m)[\alpha|\beta]| \neq 0. \tag{1.7}$$

(ii) Identify β with a subset of $N\times N$, $N = \{1,\cdots,n\}$. In the system (1.1) consider r equations given by the index set α. Let $X(\xi) = (x_{ij})_1^n$ be the unique solution of these r equations satisfying the conditions

$$x_{ij} = \xi_{\ell(i,j)},\ (i,j) \notin \beta,\quad \xi = (\xi_1,\cdots,\xi_\nu).$$

Assume that $|X(\xi)|$ does not vanish identically on $\mathbb{C}^\nu$.

Denote by $U_{\alpha,\beta}$ the set of all $(B_0,\cdots,B_m)$ which fulfill (i) and (ii). Let U be union of all $U_{\alpha,\beta}$ $\alpha \in Q_{r,(m+1)n^2}$, $\beta \in Q_{r,n^2}$ where α and β satisfying (i) and (ii). Then U is a non-empty open (algebraic) set

in $\mathcal{M}_{n,m+1}$ which contains $(A_0,\cdots,A_m)$. Moreover $(B_0,\cdots,B_m)$ is simultaneous similar to $(A_0,\cdots,A_m)$ if and only if $(B_0,\cdots,B_m) \in U$ and the following equalities hold.

$$(1.8)\qquad |L(B_0,\cdots,B_m,A_0,\cdots,A_m)[\gamma|\delta]| = 0,\quad \gamma \in Q_{r+1,(m+1)n^2},\ \delta \in Q_{r+1,n^2}.$$

Proof Let α and β satisfy the assumptions (i) and (ii). Since the coefficients of the variables $\xi_1,\cdots,\xi_\nu$ in the polynomial $|X(\xi)|$ are rational functions in the entries of $(A_0,\cdots,A_m)$ and $(B_0,\cdots.B_m)$ it follows that $U_{\alpha,\beta}$ is an open algebraic set. That is $U_{\alpha,\beta}$ is a union of a finite number of sets, each of them is characterized by some non-vanishing polynomial. Suppose that $(B_0,\cdots,B_m)$ is simultaneous similar to $(A_0,\cdots,A_m)$. Then the matrices $L(A_0,\cdots,A_m,A_0,\cdots,A_m)$ and $L(B_0,\cdots,B_m,A_0,\cdots,A_m)$ are equivalent. Therefore the above matrices have the same rank. So there exists $\alpha \in Q_{r,(m+1)n^2}$ and $\beta \in Q_{r,n^2}$ which satisfy the assumption (i). Furthermore as $B_i = TA_iT^{-1}$, $i = 0,\cdots,m$ the polynomial $|X(\xi)|$ does not vanish identically. This shows that $U_{\alpha,\beta}$ is a non-empty open algebraic set, and hence U is a non-empty algebraic open set. Moreover if $(B_0,\cdots,B_m)$ is simultaneously similar to $(A_0,\cdots,A_m)$ then $(B_0,\cdots,B_m) \in U$. In particular $(A_0,\cdots,A_m) \in U$. Assume that $(B_0,\cdots,B_m) \in U$. Then there exist a non-singular matrix X which satisfies r independent equations of (1.1). The equality (1.8) implies that all other equations of the system (1.1) are linear combinations of these r equations. So X satisfies (1.1) which means that $(B_0,\cdots,B_m)$ is simultaneous similar to $(A_0,\cdots,A_m)$. ■

Let ν and r be defined as in Theorem 1.6. Consider the following algebraic variety $\mathcal{X}$ in $\mathcal{M}_{n,m+1}$ defined by the equations

$$(1.9)\qquad |L(X_0,\cdots,X_m,A_0,\cdots,A_m)[\alpha|\beta]| = 0,\quad \alpha \in Q_{r+1,(m+1)n^2},\ \beta \in Q_{r+1,n^2}.$$

This variety splits to κ irreducible varieties

$$(1.10)\qquad \mathcal{X} = \bigcup_{i=1}^{\kappa} \mathcal{X}_i.$$

See Section 3 for various properties of algebraic varieties needed here and the appropriate references.

In Section 3 we show that κ is bounded

$$(1.11) \qquad \kappa \leq r^{\nu(mn^2+\nu)}, \quad \nu = n^2-r.$$

Let U be an open set defined in Theorem 1.6. According to Theorem 1.6

$$X \cap U \subset \mathrm{orb}(A_0,\cdots,A_m).$$

Since the orbit of $(A_0,\cdots,A_m)$ is a manifold of dimension r it follows that the point $(A_0,\cdots,A_m)$ is contained exactly in one irreducilbe variety - say X_1. Also the dimension of X_1 is r.

<u>Definition</u> 1.12. Let $M_{n,m+1,r}$ be the set of all matrices $(A_0,\cdots,A_m) \in M_{n,m+1}$ such that

$$|L(A_0,\cdots,A_m,A_0,\cdots,A_m)[\alpha|\beta]| = 0, \quad \alpha \in Q_{r+1,(m+1)n^2}, \ \beta \in Q_{r+1,n^2}.$$

By $M^0_{n,m+1,r}$ denote the open (algebraic) subset of $M_{n,m+1,r}$ of $(A_0,\cdots,A_m)$ satisfying

$$\nu(A_0,\cdots,A_m) = n^2-r .$$

(It may happen that $M^0_{n,m+1,r}$ is empty). Let $\kappa(n,m,r)$ be the maximal number of irreducible varieties X_i of dimension r in the decomposition (1.10) of the variety X given by (1.9) having a non-empty intersection with $M^0_{n,m+1,r}$ for all possible choices of $(A_0,\cdots,A_m) \in M^0_{n,m+1,r}$, ($\kappa(n,m,r) = 0$ if $M^0_{n,m+1,r}$ is empty).

So (1.11) implies

$$(1.13) \qquad \kappa(n,m,r) \leq r^{\nu(mn^2+\nu)}, \quad \nu = n^2-r.$$

<u>Theorem</u> 1.14. Let $(A_0,\cdots,A_m) \in M^0_{n,m+1,r}$. Let X_1 be the irreducible component of the variety (1.9) containing the point $(A_0,\cdots,A_m)$. Then $\mathrm{orb}(A_0,\cdots,A_m)$ is an open algebraic set in X_1. That is

$$(1.15) \qquad \mathrm{orb}(A_0,\cdots,A_m) = X_1.$$

Moreover

$$(1.16) \qquad \mathrm{orb}(A_0,\cdots,A_m) = X_1 \cap M^0_{n,m+1,r}.$$

Proof. As $\mathrm{orb}(A_0,\cdots,A_m)$ is a manifold of dimension r by Theorem 1.6 we get $\mathrm{orb}(A_0,\cdots,A_m) \subset X_1$. Let U be defined as in Theorem 1.6. According to Theorem 1.6

$$\mathrm{orb}(A_0,\cdots,A_m) = X_1 \cap U.$$

As U is an open algebraic set in $M_{n,m+1}$ we get that $\mathrm{orb}(A_0,\cdots,A_m)$ is an open algebraic set in the irreducible variety X_1. So (1.15) holds. (Here $\overline{\mathrm{orb}}(A_0,\cdots,A_m)$ means the closure of the orbit set of $(A_0,\cdots,A_m)$). Let $(B_0,\cdots,B_m) \in X_1 \cap M^0_{n,m+1,r}$. So any neighborhood of $(B_0,\cdots,B_m)$ contains a point $(C_0,\cdots,C_m) \in \mathrm{orb}(A_0,\cdots,A_m)$. According to the equalities (1.3) the matrices $L(B_0,\cdots,B_m,A_0,\cdots,A_m)$ and $L(B_0,\cdots,B_m,C_0,\cdots,C_m)$ are equivalent. As $(B_0,\cdots,B_m) \in X_1$ we deduce

$$|L(B_0,\cdots,B_m,C_0,\cdots,C_m)[\alpha|\beta]| = 0, \quad \alpha \in Q_{r+1,(m+1)n^2}, \ \beta \in Q_{r+1,n^2}.$$

As analogous results to Theorem 1.6 yields that there exists a neighborhood $\hat{U}$ of $(B_0,\cdots,B_m)$ such that $(C_0,\cdots,C_m) \in \mathrm{orb}(B_0,\cdots,B_m)$ provided that $(C_0,\cdots,C_m) \in \hat{U}$ and the above equalities hold. So there exist $(C_0,\cdots,C_m)$ which lies in the orbits generated by $(A_0,\cdots,A_m)$ and $(B_0,\cdots,B_m)$. Whence $(B_0,\cdots,B_m) \in \mathrm{orb}(A_0,\cdots,A_m)$. As $\mathrm{orb}(A_0,\cdots,A_m) \subset M^0_{n,m+1,r}$ we finally deduce the equality (1.16). []

Consider the left hand side of the equalities (1.9). These are multilinear polynomials in the entries of $X_0,\cdots,X_m$. Let $\mathcal{W}_{n,m+1}$ be the linear space of all multilinear polynomials in the entries of $X_0,\cdots,X_m$. We formally define $\mathcal{W}_{n,m+1}$ as follows. Denote by $M_n^{(0,1)}$ the set of all $n \times n$ matrices $\gamma = (\gamma_{ij})$ with $(0,1)$ entries, i.e. $\gamma_{ij} = 0,1$. For $X = (x_{ij}) \in M_n$ we let

$$X^\gamma = \prod_{1 \le i,j \le n} x_{ij}^{\gamma_{ij}}, \ \gamma \in M_n^{(0,1)} \quad (x^0_{ij} \equiv 1).$$

Then $\mathcal{W}_{n,m+1}$ consists of all polynomials

$$(1.17) \qquad P(X_0,\cdots,X_m) = \sum_{\gamma_i \in M_n^{(0,1)}} p_{\gamma_0,\cdots,\gamma_m} X_0^{\gamma_0},\cdots,X_m^{\gamma_m}.$$

Clearly

$$(1.18) \qquad \dim \mathcal{W}_{n,m+1} = 2^{(m+1)n^2}.$$

With this notation

$$|L(X_0,\cdots,X_m,A_0,\cdots,A_m)[\alpha|\beta]| =$$

$$(1.19) \qquad \sum_{\gamma_i \in M_n^{(0,1)}} P_{\gamma_0,\cdots,\gamma_m}(A_0,\cdots,A_m,\alpha,\beta)X_0^{\gamma_0},\cdots,X_m^{\gamma_m},$$

$$\alpha \in Q_{r+1,(m+1)n^2}, \quad \beta \in Q_{r+1,n^2}.$$

Denote by $P(A_0,\cdots,A_m)$ the

$$\binom{(m+1)n^2}{r+1}\binom{n^2}{r+1} \times 2^{(m+1)n^2}$$

matrix whose rows are the vectors

$$(P_{\gamma_0,\cdots,\gamma_m}(A_0,\cdots,A_m,\alpha,\beta)), \quad \gamma_0,\cdots,\gamma_m \in M_n^{(0,1)}.$$

In fact the matrix $P(A_0,\cdots,A_m)$ contains a lot of zero columns, Indeed if we pick up a minor of the form (1.19) then each entry of such a minor consists of a sum one element of some X_k and some A_j. Therefore the polynomial (1.19) contains exactly $(r+1)!$ moninomials of degree $r+1$ of the form $\xi_1,\cdots,\xi_{r+1}$ in the entries of $X_0,\cdots,X_m$. So the number of distinct mononomials of degree $r+1$ appearing in all possible expression in (1.19) is exactly $(r+1)!\binom{(m+1)n^2}{r+1}\binom{n^2}{r+1}$. The number of distinct monomials of degree $d \leq r$ which may appear in (1.19) is $d!\binom{(m+1)n^2}{d}\binom{n^2}{d}$. Hence the upper bound for nontrivial columns appearing in $P(A_0,\cdots,A_m)$ is

$$(1.20) \qquad N_{n,m,r} = \sum_{d=0}^{r+1} d!\binom{(m+1)n^2}{d}\binom{n^2}{d} \leq 2^{r+1}\binom{(m+1)n^2}{r+1}\binom{n^2}{r+1}.$$

By $\mathcal{M}_{n,m+1,r,\rho}$ denote the subset of all tuples $(A_0,\cdots,A_m) \in \mathcal{M}_{n,m+1,r}$ such that rank $P(A_0,\cdots,A_m) \leq \rho$.

Clearly, $M_{n,m+1,r,\rho}$ is an algebraic set. Let $M^0_{n,m+1,r,\rho}$ be the subset of all tuples $(A_0,\cdots,A_m) \in M^0_{n,m+1,r}$ satisfying

$$\text{rank } P(A_0,\cdots,A_m) = \rho. \tag{1.21}$$

Again, it is easy to see that $M^0_{n,m+1,r,\rho}$ is an open (algebraic) set in $M_{n,m+1,r,\rho}$. We now give sets of invariant rational functions which determine uniquely a finite number of orbits lying in $M^0_{n,m+1,r,\rho}$. To do that we need to recall the notion of Hermite normal form of a rectangular matrix A. See for example Marcus-Minc [1964]. Two $p\times q$ rectangular matrices A and B are called row equivalent $(A\sim B)$ if there exists a nonsingular matrix Q such that $B=QA$. Any $p\times q$ matrix A can be brought to the unique Hermite normal form $E = E(A)$ using the elementary row operations. $E = (e_{ij})$ is characterized by the conditions

$$\begin{gathered} 1 \leq p_1 < p_2 < \cdots < p_\rho \leq p, \quad \rho = \text{rank } A, \\ e_{ip_i} = 1, \quad e_{ij} = 0 \quad \text{for } j < p_i, \quad \text{for } j = p_{i+1},\cdots,p_\rho, \quad i=1,\cdots,\rho \\ e_{iq} = 0 \quad \text{for } i > \rho. \end{gathered} \tag{1.22}$$

The integers $p_1,\cdots,p_\rho$ are called the discrete invariants of A and the entries e_{iq}, $p_i < q \neq p_{i+1},\cdots,p_\rho$, $i = 1,\cdots,\rho$, are called the continuous invariants of A. Given $p_1,\cdots,p_\rho$ then the continuous invariants are well determined rational functions of A.

<u>Theorem</u> 1.23. Let r and ρ be positive integers such that $M^0_{n,m+1,r,\rho}$ is nonempty. Assume that $(A_0,\cdots,A_m),(B_0,\cdots,B_m) \in M^0_{n,m+1,r,\rho}$. If $(A_0,\cdots,A_m)$ and $(B_0,\cdots,B_m)$ are simultaneously similar then

$$P(A_0,\cdots,A_m) \sim P(B_0,\cdots,B_m). \tag{1.24}$$

Moreover, there are at most $\kappa(n,m,r)$ distinct orbits in $M^0_{n,m+1,r,\rho}$ which satisfy the equality (1.24). That is the discrete and the continuous invariants of $P(A_0,\cdots,A_m)$ determine at most $\kappa(n,m,r)$ distinct orbits in $M^0_{n,m+1,r,\rho}$.

Proof. Assume that $(B_0,\cdots,B_m) = T(A_0,\cdots,A_m)T^{-1}$. Then (1.4) yields that the matrices $L(X_0,\cdots,X_m,A_0,\cdots,A_m)$ and $L(X_0,\cdots,X_m,B_0,\cdots,B_m)$ are equivalent. Moreover the transformation matrices do not depend on the matrices $X_0,\cdots,X_m$. So any $k\times k$ minor of $L(X_0,\cdots,X_m,B_0,\cdots,B_m)$ $(L(X_0,\cdots,X_m,A_0,\cdots,A_m))$ is a fixed linear combination of all $k\times k$ minors of $L(X_0,\cdots,X_m,A_0,\cdots A_m)$ $(L(X_0,\cdots,X_m,B_0,\cdots,B_m))$. So the matrices $P(A_0,\cdots,A_m)$ and $P(B_0,\cdots,B_m)$ can be transformed each one to the other one by elementary row operations. That is the equality (1.24) holds.

Consider the algebraic variety defined by (1.9). In the decomposition (1.10) we may assume that $X_1,\cdots,X_\tau$ are all the irreducible varieties which have exactly the dimension r such that $X_i \cap M^0_{n,m+1,r,\rho}$ is an open non-empty set in X_i for $i = 1,\cdots,\tau$. By the definition of $\kappa(n,m,r)$, $\tau \leq \kappa(n,m,r)$. Suppose that

$$(A_0^{(i)},\cdots,A_m^{(i)}) \in X_i \cap M^0_{n,m+1,r,\rho}, \quad i = 1,\cdots,\tau$$

and assume that $(A_0^{(i)},\cdots,A_m^{(i)})$ is a regular point in X_i. That is X_i is a manifold of dimension r in the neighborhood of $(A_0^{(i)},\cdots,A_m^{(i)})$. We assume the normalization $A_k^{(1)} = A_k$, $k = 0,\cdots,m$. So $X_1 = \overline{\mathrm{orb}}(A_0,\cdots,A_m)$. The equalities (1.4) yield that $\mathrm{orb}(A_0^{(i)},\cdots,A_m^{(i)})$ satisfies the system (1.9). As $\mathrm{orb}(A_0^{(i)},\cdots,A_m^{(i)})$ is a manifold of dimension r which passes through a regular point $(A_0^{(i)},\cdots,A_m^{(i)})$ of X_i we deduce that $(A_0^{(i)},\cdots,A_m^{(i)}) \subseteq X_i$. Now Theorem 1.14 implies that $\overline{\mathrm{orb}}(A_0^{(i)},\cdots,A_m^{(i)})$ is an irreducible variety. So $\overline{\mathrm{orb}}(A_0^{(i)},\cdots,A_m^{(i)}) = X_i$. Let $(B_0,\cdots,B_m) \in M^0_{n,m+1,r,\rho}$ and assume that (1.24) holds. Thus any $(r+1)\times(r+1)$ minor of $L(X_0,\cdots,X_m,A_0,\cdots A_m)$ is a linear combination of all $k\times k$ minors of $L(X_0,\cdots,X_m,B_0,\cdots,B_m)$ and vice versa. Therefore $(X_0,\cdots,X_m)$ satisfies the system (1.9) if and only if $(X_0,\cdots,X_m)$ satisfies the system

(1.25) $$L(X_0,\cdots,X_m,B_0,\cdots,B_m)[\alpha|\beta] = 0, \quad \alpha \in Q_{r+1,(m+1)n^2}, \ \beta \in Q_{r+1,n^2}.$$

As $\nu(B_0,\cdots,B_m) = n^2-r$, $(B_0,\cdots,B_m)$ is a solution to the above

equalities. Whence $(B_0,\cdots,B_m)$ satisfies the equations (1.9). Therefore $\mathrm{orb}(B_0,\cdots B_m)$ satisfies the equalities (1.9). Since $(B_0,\cdots.B_m) \in M^0_{n,m+1,r,\rho}$, we deduce that $\mathrm{orb}(B_0,\cdots,B_m) \subset M^0_{n,m+1,r,\rho}$. So $\overline{\mathrm{orb}}(B_0,\cdots,B_m)$ is an irreducible variety of dimension r such that its intersection with ${}^0_{n,m+1,r,\rho}$ is a non-empty open set in $\overline{\mathrm{orb}}(B_0,\cdots,B_m)$. Hence $\overline{\mathrm{orb}}(B_0,\cdots,B_m) = {}_i$ for some $1 \leq i \leq \tau$. As $(B_0,\cdots,B_m) \in M^0_{n,m+1,r}$ Theorem 1.14 yields that $(B_0,\cdots,B_m)$ is simultaneously similar to $(A_0^{(i)},\cdots, A_m^{(i)})$. Thus, then at most τ $(\leq \kappa(n,m,r))$ distinct orbits such that any two points on those orbits satisfy the equalities (1.24). []

Let $(A_0,\cdots,A_m) \in M^0_{n,m+1,r,\rho}$. Assume that $\omega = (p_1,\cdots,p_\rho)$ is the set of discrete invaviants of $P(A_0,\cdots,A_m)$. Let $V^0(\omega)$ be of the set of all $(B_0,\cdots,B_m) \in M^0_{n,m+1,r,\rho}$ for which ω is the set of the discrete invariants of $P(B_0,\cdots,B_m)$. Clearly $V^0(\omega)$ is an open set in $M^0_{n,m+1,r,\rho}$. Moreover, $M^0_{n,m+1,r,\rho}$ is a finite union of set $V^0(\omega)$, for admissible $\omega \in Q_{\rho,n_2}$, $n_2 = N_{n,m,r}$. On each $V^0(\omega)$ the set of continuous invariants of $P(A_0,\cdots,A_m)$ - which are rational functions in the entries of $A_0,\cdots,A_m$ - classify the orbits in $V^0(\omega)$ up to a finite number not exceeding $\kappa(n,m,r)$. We conjecture

<u>Conjecture</u> 1.26. Let $(A_0,\cdots,A_m)$, $(B_0,\cdots,B_m) \in M^0_{n,m+1,r,\rho}$. Assume that the equality (1.24) holds. Then $(A_0,\cdots,A_m)$ and $(B_0,\cdots,B_m)$ are simultaneously similar.

For $m = 0$ the above conjecture is valid. Indeed, according to Friedland [1980]

$$(1.27) \qquad \nu(B_0,A_0) \leq [\nu(A_0,A_0) + \lambda(B_0,B_0)]/2$$

and the equality sign holding if and only if A_0 and B_0 are similar. According to the proof of Theorem 1.23 the equality (1.24) implies that all $(r+1)\times(r+1)$ minors of $L(B_0,A_0)$ vanish. So

$$\nu(B_0,A_0) \geq n^2-r = \nu(A_0,A_0) = \nu(B_0,B_0).$$

Hence we must have the equality sign in (1.27) which means that A_0 and B_0 are similar. In the next section we verify the above conjecture for $m=1$ and $n=2$.

2. The 2×2 case.

Let $A, B, \in M_2$. Clearly

(2.1) $\phi_1 = tr(A)$, $\phi_2 = tr(A^2)$, $\phi_3 = tr(B)$, $\phi_4 = tr(B^2)$, $\phi_5 = tr(AB)$

are invariant polynomials under the simultaneous similarity.

Theorem 2.2. Let U be the following algebraic variety

(2.3) $U = \{(A,B) | \{2tr(A^2)-[tr(A)]^2\}\{2tr(B^2)-[tr(B)]^2\} = [2tr(AB) - tr(A)tr(B)]^2\}$

Then, for any pair (A,B) not lying in U the orbit of (A,B) is determined uniquely by the values of $\phi_i(A,B)$, $i = 1,\cdots,5$.

Proof. Suppose first that A has distinct eigenvalues λ_1, λ_2, $\lambda_1 \neq \lambda_2$. That is $[tr(A)]^2 \neq 2tr(A^2)$. The eigenvalues λ_1 and λ_2 are determined by the values of $tr(A)$ and $tr(A^2)$. Then we can choose a pair (D,E) lying in the orbit of (A,B) such that

(2.4) $$D = D(\lambda_1,\lambda_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad E = E(e_{11},e_{22},e_{12},e_{21}) = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}.$$

Now

$$tr(E) = e_{11}+e_{22}, \quad tr(DE) = \lambda_1 e_{11}+\lambda_2 e_{22}, \quad tr(E^2) = e_{11}^2+e_{22}^2+2e_{12}e_{21}.$$

Thus

(2.5) $$e_{11} = \frac{tr(DE) - \lambda_2 tr(E)}{\lambda_1 - \lambda_2}, \qquad e_{22} = \frac{tr(DE) - \lambda_1 tr(E)}{\lambda_2 - \lambda_1}$$

$$e_{12}e_{21} = [tr(E^2) - e_{11}^2 - e_{22}^2]/2$$

We claim that $e_{11}e_{22} \neq 0$. Otherwise $(D,E) \in U$, i.e. $(A,B) \in U$ contrary to our assumption. Therefore by considering XDX^{-1} and XEX^{-1}, where X is a diagonal matrix, we may assume

(2.6) $$e_{21} = 1.$$

In that case D and E are determined uniquely by

$tr(D) = tr(A)$, $tr(D^2) = tr(A^2)$, $tr(E) = tr(B)$,

$tr(E^2) = tr(B^2)$, $tr(DE) = tr(AB)$.

Suppose that $[tr(A)]^2 = 2tr(A^2)$. Consider the matrix $\alpha A+\beta B$. Then

$$2tr[(\alpha A+\beta B)^2] - [tr(\alpha A+\beta B]^2 =$$

$$\alpha^2\{2tr(A^2) - [tr(A)]^2\} + B^2\{2tr(B^2) - [tr(B)]^2\} \tag{2.7}$$

$$+ 2\alpha\beta\{2tr(AB) - tr(A)tr(B)\}.$$

Thus $\alpha A+\beta B$ has a double eigenvalue for all α and β if and only if A,B) lies on the variety

$$V = \{(A,B)|2tr(A^2) = [tr(A)]^2,\ 2tr(B^2) = [tr(B)]^2,\ 2tr(AB) = tr(A)tr(B)\}. \tag{2.8}$$

Since $V \subset U$ we see that $(A,B) \notin V$. So we can choose $\beta \neq 0$ such that A_1 $A + \beta B$ has distinct eigenvalues. Thus, there exists a matrix X such that

$$XA_1X^{-1} = D, \quad XBX^{-1} = E.$$

Again, if $e_{11}e_{22} = 0$ then A_1 and B are simultaneously similar to upper (lower) triangular matrices. So A and B are simultaneously similar to upper (lower) triangular matrices which contradicts the assumption $(A,B) \in U$. That is the orbit of (A,B) contains a matrix of the form $D-\beta E$ and E where D is diagonal, $e_{21} = 1$ and D and E are defined uniquely by $_i(A,B)$, $i = 1,\cdots,5$ having fixed the value of β. ■

We now examine the matrix meaning of the variety U.

<u>Theorem</u> 2.9. A pair of matrices (A,B) belongs to U if and only if (A,B) is simultaneously similar to a pair of upper triangular matrices.

<u>Proof.</u> Assume first that A has two distinct eigenvalues λ_1,λ_2, $\lambda_1 \neq \lambda_2$. Then (A,B) is simultaneously similar to a pair (D,E) where D and E are given by (2.4). A straightforward calculation yields

$$2tr(A^2) - [tr(A)]^2 = 2tr(D^2) - [tr(D)]^2 = (\lambda_1-\lambda_2)^2$$

$$2tr(B^2) - [tr(B)]^2 = 2tr(E^2) - [tr(E)]^2 = (e_{11}-e_{22})^2 + 4e_{12}e_{21} \tag{2.10}$$

$$2tr(AB) - tr(A)tr(B) = 2tr(DE) - tr(D)tr(E) = (\lambda_1-\lambda_2)(e_{11}-e_{22}).$$

The assumption that $(A,B) \in U$ means

$$(\lambda_1-\lambda_2)^2[(e_{11}-e_{22})^2+4e_{12}e_{21}] = (\lambda_1-\lambda_2)^2(e_{11}-e_{22})^2.$$

Since $\lambda_1 \neq \lambda_2$ we get that $e_{12}e_{21} = 0$. If $e_{21} = 0$ then D and E are upper triangular. Suppose that $e_{12} = 0$. Define

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then (A,B) are simultaneously similar to the upper diagonal matrices (PDP^{-1}, PEP^{-1}). This establishes the lemma in case that A has distinct eigenvalues. Assume that A has a double eigenvalue. If $A = \lambda I$ then clearly (A,B) are a simultaneously similar pair of upper diagonal matrices. So suppose that (A,B) are simultaneously similar to a pair (D,E)

$$(2.11) \qquad D = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad E = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}.$$

Since $(A,B) \in U$ we must have

$$0 = 2\mathrm{tr}(AB) - \mathrm{tr}(A)\mathrm{tr}(B) = 2\mathrm{tr}(DE) - \mathrm{tr}(D)\mathrm{tr}(E) = 2e_{21}.$$

that is (D,E) are upper triangular. ∎

<u>Corollary</u> 2.12. The variety U is invariant under the linear transformation

$$(2.13) \qquad A_1 = \alpha A + \beta B, \quad B_1 = \gamma A + \delta B, \quad \alpha\delta - \beta\gamma \neq 0.$$

Let

$$(2.14) \qquad \begin{aligned} U_1 &= \{(A,B) \mid (A,B) \in U, \ \ 2\mathrm{tr}(A^2) \neq [\mathrm{tr}(A)]^2\} \\ U_2 &= \{(A,B) \mid (A,B) \in U, \ \ 2\mathrm{tr}(B^2) \neq [\mathrm{tr}(B)^2]\}. \end{aligned}$$

<u>Theorem</u> 2.15. On U_1 or U_2 the values of $\phi_i(A,B)$, $i = 1,\cdots,5$ correspond to 3 distinct orbits. On $U_1 \cap U_2$ the values of $\phi_i(A,B)$, $i = 1,2,3,4$ correspond to 6 distinct orbits, while on $U_1 \cup U_2 - U_1 \cap U_2$ these values correspond to 3 distinct orbits.

<u>Proof:</u> Suppose that $(A,B) \in U_1$. Then (A,B) is simultaneous similar to (D,E) given by (2.4). From the proof of Theorem 2.9 it follows that $e_{12}e_{21} = 0$. So we have the possibilities $e_{12} = 0$ and $e_{21} \neq 0$, $e_{21} = 0$ and $e_{12} \neq 0$ and $e_{12} = e_{21} = 0$. As in the proof of Theorem 2.2 using diagonal similarity we can assume that E may have the following form

$$E_1 = E(e_{11},e_{22},1,0), \ E_2 = E(e_{11},e_{22},0,1), \ E_3 = E(e_{11},e_{22},0,0).$$

We claim that the pairs (D,E_i) and $D,E_j)$ are not simultaneously similar for $i\neq j$. Indeed, suppose

$$D = XDX^{-1}, \quad E_j = XEX^{-1}, \quad i\neq j.$$

Since D is diagonal with distinct eigenvalues the first equality implies that X is also diagonal. Then the second equality is impossible. Let μ_1 and μ_2 be the eigenvalues of B. Clearly $\{\mu_1,\mu_2\} = \{e_{11},e_{22}\}$. So either $\mu_1 = e_{11}$, $\mu_2 = e_{22}$ or $\mu_1 = e_{22}$, $\mu_2 = e_{11}$. Let

$$\bar{E}_1 = E(e_{22},e_{11}1,0), \ \bar{E}_2 = E(e_{22},e_{11},0,1), \ \bar{E}_3 = E(e_{22},e_{11},0,0).$$

If $\mu_1 \neq \mu_2$ then $\bar{E}_j \neq XE_iX^{-1}$ for any non-singular diagonal matrix X. This shows that for $(A,B) \in U_1 \cap U_2$ the values of the functions $\phi_i(A,B)$, $i = 1,2,3,4$ corresponds to 6 distinct orbits. If B has a multiple eigenvalue then $\bar{E}_i = E_i$ and the above values correspond to 3 distinct orbits. Thus we proved that on $U_1 \cup U_2 - U_1 \cap U_2$ the values of $\phi_i(A,B)$, $1 \leq i \leq 4$, determine 3 distinct orbits. Suppose we are also given $\phi_5(A,B) = tr(AB)$. Note that

$$(\lambda_1\mu_1+\lambda_2\mu_2) - (\lambda_1\mu_2+\lambda_2\mu_1) = (\lambda_1-\lambda_2)(\mu_1-\mu_2).$$

Thus for $(A,B) \in U_1 \cap U_2$

$$tr(DE_i) \neq tr(D\bar{E}_j).$$

Therefore on $U_1 \cup U_2$ the values of $\phi_1(A,B)$, $1 \leq i \leq 5$ correspond to 3 distinct orbits. ■

Let $(A,B) \in U$. Assume that (A,B) is simultaneously similar to diagonal matrices (D,E_3). Then A and B commutes. That is (A,B) lie on the manifold

$$C = \{(A,B)\,|\,AB-BA = 0\}. \tag{2.16}$$

Clearly $U \cap C$ is a subvariety of U. That is a generic orbit in U contains a pair of the form (D,E_1). The fact that the orbit corresponding to (D,E_3) is not generic can be seen in the following way. Consider the orbit $orb(A,B)$ as a manifold. Denote by $\dim orb(A,B)$ the dimension of

this manifold. Then

$$\dim \operatorname{orb}(A,E_i) = 3, \qquad A \in U_1,$$

(2.17)

$$\dim \operatorname{orb}(D,E_3) = 2, \qquad D \in U_1$$

Indeed consider all matrices X which commutes with D and E_i

$$DX - XD = 0, \quad E_iX - XE_i = 0, \quad D \in U_1.$$

For $i=1,2$ $X = \lambda I$ and for $i=3$ X is any diagonal matrix. This establishes (2.17). We shall see later in this section that $\phi_i(A,B)$, $1 \leq i \leq 5$ are generators of $[V]^G$.

From the proof of Theorem 2.15 it follows that the values of these functions do not separate between the two generic orbits corresponding to the pairs (D,E_1) and (D,E_2), $D \in U_1$. According to the result in Friedland [1981] there exist rational functions $\theta_i \in (U)^G$, $i = 1,\cdots,r$ which separate the orbits in $U-W$, for some algebraic subvariety W in U. We now give such functions θ_i. Let $(A,B) \in U$. Then Theorem 2.9 claims that A and B simultaneously similar to a pair of upper triangular matrices. Suppose that $(A,B) \in U_1 \cap U_2$ and $\dim \operatorname{orb}(A,B) = 3$. Then $A = (a_{ij})_1^2$ and $B = (b_{ij})_1^2$ have exactly one common eigenvector $(x_1,x_2)^t$. Corresponding to the eigenvalues λ_1 and μ_1 respectively

$$(a_{11}-\lambda_1)x_1 + a_{12}x_2 = 0 \qquad a_{21}x_1 + (a_{22}-\lambda_1)x_2 = 0$$

$$(b_{11}-\mu_1)x_1 + b_{12}x_2 = 0 \qquad b_{21}x_1 + (b_{22}-\mu_1)x_2 = 0.$$

Assume for simplicity that

$$a_{12}a_{21}b_{12}b_{21}x_1x_2 \neq 0.$$

Then we have the equalities

$$(a_{11}-\lambda_1)b_{12} = (b_{11}-\mu_1)a_{12}, \quad (a_{22}-\lambda_1)b_{21} = (b_{22}-\mu_1)a_{21}.$$

So

$$\mu_1 = \frac{b_{12}}{a_{12}}\lambda_1 + \frac{b_{11}a_{12}-a_{11}b_{12}}{a_{12}} = \frac{b_{21}}{a_{21}}\lambda_1 + \frac{b_{22}a_{21}-a_{22}b_{21}}{a_{21}}.$$

From the last two equalities we can compute the value of λ_1 and then the value of μ_1

$$\lambda_1 = \alpha(A,B) = \frac{(b_{11}-b_{22})a_{12}a_{21} + a_{22}a_{12}b_{21} - a_{11}a_{21}b_{12}}{a_{12}b_{21} - a_{21}b_{12}}$$

(2.18)

$$\mu_1 = \beta(A,B) = \frac{(a_{11}-a_{22})b_{12}b_{21} + b_{22}b_{12}a_{21} - b_{11}b_{21}a_{12}}{b_{12}a_{21} - b_{21}a_{12}}$$

Theorem (2.19). The functions $\alpha(A,B)$, $\beta(A,B)$ belong to $(U)^G$. Moreover these functions are defined on $\mathrm{orb}(A,B)$ such that $(A,B) \in U-C$ and the values of $\alpha(A,B)$, $\beta(A,B)$, $\mathrm{tr}(A)$, $\mathrm{tr}(B)$ determine these orbits uniquely.

Proof. Clearly, $\alpha(A,B)$, $\beta(A,B) \in (U)$ since $a_{12}b_{21} - a_{21}b_{12}$ is not vanishing identically on U. Put

$$\lambda_2 = \mathrm{tr}(A) - \alpha(A,B), \quad \mu_2 = \mathrm{tr}(B) - \beta(A,B).$$

Then a straightforward calculation yields

$$\lambda_1^2 + \lambda_2^2 = \mathrm{tr}(A^2), \quad \mu_1^2 + \mu_2^2 = \mathrm{tr}(B^2),$$

as $(A,B) \in U$. That is $\alpha(A,B)$ and $\beta(A,B)$ are the eigenvalues of A and B respectively. This shows that $\alpha,\beta \in (V)^G$. Assume that $A \in U_1$ and $AB-BA \neq 0$. Then the orbit of (A,B) contains a pair (D,E_1). Clearly $\alpha(A,B)$ and $\beta(A,B)$ is not defined for (D,E_1). Let $(P,Q) \in \mathrm{orb}(A,B)$ lying closely to (D,E_1). That is

$$P = (I+X)^{-1}D(I+X), \quad Q = (I+X)^{-1}E_1(I+X),$$
$$P = (p_{ij})_1^2, \quad Q = (q_{ij})_1^2 .$$

Then

$$p_{12} = (\lambda_1-\lambda_2)x_{12} + O(\|X\|^2), \; p_{21} = (\lambda_2-\lambda_1)x_{21} + O(\|X\|^2)$$
$$q_{12} = 1 + O(\|X\|), \quad q_{21} = (\mu_2-\mu_1)x_{21} + O(\|X\|^2).$$

So

$$p_{12}q_{21}-p_{21}q_{12} = (\lambda_1-\lambda_2)(\mu_2-\mu_1)x_{12}x_{21} - (\lambda_2-\lambda_1)x_{21} + O(\|X\|^2).$$

Whence it is possible to find $(P,Q) \in \mathrm{orb}(A,B)$ such that the above expression is different from zero. Put

$$\alpha(A,B) = \alpha(P,Q), \quad \beta(A,B) = \beta(P,Q)$$

and the functions α and β are well defined. The matrices (D,E_1) have exactly one common eigenvalue which corresponds to λ_1 and e_{11}. So

$$\lambda_1 = \alpha(A,B), \qquad e_{11} = \beta(A,B)$$

$$\lambda_2 = tr(A) - \alpha(A,B), \quad e_{22} = tr(B) - \beta(A,B)$$

and the matrices D and E_1 are determined. The same arguments apply if $B \in U_2$, $AB-BA \neq 0$. Suppose that $A \notin U_1$ and $B \notin U_2$. That is (A,B) is simultaneously similar to $(E(\lambda,\xi), E(\mu,\eta))$ where

(2.20) $$E(\lambda,\xi) = \begin{pmatrix} \lambda & \xi \\ 0 & \lambda \end{pmatrix}.$$

But then $E(\lambda,\xi)$ and $E(\mu,\eta)$ commute. The proof of the theorem is completed. ∎

For any 2×2 matrices A,B the $(1,1)$ entry of $AB-BA$ equals to $a_{12}b_{21} - a_{21}b_{12}$. That is $a_{12}b_{21} - a_{21}b_{12}$ vanishes identically on C and therefore the functions $\alpha(A,B)$ and $\beta(A,B)$ are not defined on C.

If $A \in U_1$ and A and B commutes so A and B are simultaneously diagonal then from the proof of Theorem 2.15 it follows

<u>Theorem</u> 2.21. On $(U_1 \cup U_2 \cap C)$ the values of the functions $\phi_i(A,B)$, $i = 1,2,3,4,5$ determine the orbit of (A,B).

We are left with the orbits in $U \cap C$ such that A and B have double eigenvalues. It means that (A,B) are simultaneously similar to $(E(\lambda,\xi), E(\mu,\eta))$. As $E(\lambda,\xi)$ and $E(\mu,\eta)$ commutes we get

(2.22) $$V \subseteq C.$$

Thus we need to classify the orbits in V.

<u>Theorem</u> 2.23. Let $(A,B) \in V$. Then

(2.24) $$\gamma(A,B) = a_{12}/b_{12} = a_{21}/b_{21}$$

belongs to $(V)^G$. Suppose that either $\gamma(A,B)$ or $1/\gamma(A,B)$ is defined on the orbit of (A,B). Then the values of $tr(A), tr(B)$ and $\gamma(A,B)(1/\gamma(A,B))$ determine a unique orbit in V.

Proof. According to what we proved $\operatorname{orb}(A,B)$ contains a pair $(E(\lambda,\xi), E(\mu,\eta))$. Clearly

$$\lambda = \tfrac{1}{2}\operatorname{tr}(A), \quad \mu = \tfrac{1}{2}\operatorname{tr}(B).$$

By considering the matrices

$$A_1 = A - \lambda I, \quad B_1 = B - \mu I$$

we can assume that $\lambda=0$ and $\mu=0$. Let

$$A = XE(0,\xi)X^{-1}, \qquad B = XE(0,\eta)X^{-1}.$$

Then a straightforward calculation shows that $\gamma(A,B) = \xi/\eta$.

That is $\gamma(A,B)$ is indeed an invariant function on V. Assume that $\eta \neq 0$. In that case $\gamma(A,B)$ is well defined. Choosing an appropriate diagonal matrix X and considering the matrices $XE(\lambda,\xi)X^{-1}$, $XE(\mu,\eta)X^{-1}$ we can assume that $\eta=1$. Then $\xi = \gamma(A,B)$. Hence, $\operatorname{tr}(A)$, $\operatorname{tr}(B)$ and $\gamma(A,B)$ determine the orbit of (A,B). The same arguments apply if $\xi \neq 0$, i.e. $1/\gamma(A,B)$ is well defined. ■

Suppose that neither $\gamma(A,B)$ nor $1/\gamma(A,B)$ are not defined on the orbit of (A,B). Then we must have that (A,B) lie on W

$$W = \{(A,B) \mid a_{12} = a_{21} = b_{21} = b_{12} = 0, \ a_{11} = a_{22}, \ b_{11} = b_{22}\}. \tag{2.25}$$

Clearly, on W $\operatorname{tr}(A)$ and $\operatorname{tr}(B)$ determine the orbit completely.

Thus we completed the classification of the simultaneous similarity of 2×2 pairs of matrices according to the program outlined in Friedland [1981]. Next we note that on all subvarieties in $M_2 \times M_2$ except V the values of the functions $\phi_i(A,B)$, $1 \leq i \leq 5$ determine a finite number of orbits. We claim that on V the values of any set of invariant polynomial functions, i.e. functions belonging to $[V]^G$, cannot in general determine a finite number of orbits in V.

Theorem 2.26. The functions $\operatorname{tr}(A)$ and $\operatorname{tr}(B)$ are the generators in $[V]^G$.

Proof. Let $(A,B) \in V$. Then $\mathrm{orb}(A,B)$ contains a pair $E(\lambda,\xi)$, $E(\mu,\eta)$. Then

$$E^{-1}(1,d,0,0)\,E(\lambda,\xi)E(1,d,0,0) = E(\lambda,\xi d)$$

$$E^{-1}(1,d,0,0)E(\mu,\eta)E(1,d,0,0) = E(\mu,\eta d)$$

letting $d \to 0$ we get that the closure of $\mathrm{orb}(A,B)$ contains the matrices $E(\lambda,0)$, $E(\mu,0)$. Suppose that $\phi(A,B) \in [V]$. Then

$$\phi(A,B) = \phi(E(\lambda,0),E(M,0)) = g(\lambda,\mu) = h(\mathrm{tr}(A),\mathrm{tr}(B)),$$

for some $h \in C[x,y]$. ∎

Thus the transcendence degree of $[V]^G$ over C is 2. The transcendence degree of $(V)^G$ over C is 3. More precisely we have

Theorem 2.27. The functions $\mathrm{tr}(A)$, $\mathrm{tr}(B)$ and $\gamma(A,B)$ generate $(V)^G$ and they are algebraically independent.

Proof We claim that

(2.28) $\dim \mathrm{orb}(A,B) \leq 2, \quad (A,B) \in V.$

Indeed, suppose that (A,B) is simultaneously similar to $E(\lambda,\xi)$, $E(\mu,\eta)$. Then $E(\alpha,\beta)$ commutes always with $E(\lambda,\xi)$ and $E(\mu,\eta)$. Moreover if either ξ or $\eta \neq 0$ then $E(\alpha,\beta)$ is the only matrix which commute with $E(\lambda,\xi)$ and $E(\mu,\eta)$. This proves (2.28). As the dimension of V is 5 the transcendence degree of $(V)^G$ is at most 3. Clearly $\mathrm{tr}(A)$, $\mathrm{tr}(B)$ and $\gamma(A,B)$ can be given any values z_1, z_2 and z_3. So these functions are algebraically independent, i.e. the transcendence degree of $(V)^G$ is 3. Let $\phi(A,B) \in (V)$. Then $\phi(A,B)$ is algebraic with respect to $x_1 = \mathrm{tr}(A)$, $x_2 = \mathrm{tr}(B)$, $x_3 = \gamma(A,B)$

(2.29) $$\phi^m + \sum_{i=1}^{m} \rho_i(x_1,x_2,x_3)\phi^{m-i} = 0, \qquad \rho_i \in C(x_1,x_2,x_3)$$

be the minimal equation for ϕ. So the left hand side of (2.29) is an irreducible polynomial over $C(x_1,x_2,x_3)$. But then given the values of x_1,x_2,x_3 we know that $\mathrm{orb}(A,B)$ is determined (x_3 is well defined).

Thus, if ϕ is defined on orb(A,B), for example $\rho_i(x_1,x_2,x_3) \neq 0$, then ϕ has unique value. But it is well known that (2.29) must be m sheeted cover of **C**. Whence $m=1$ and ϕ is rational in x_i, $i = 1,2,3$. ∎

Next we use the arguments of Theorem (2.27) to show that $[M_2 \times M_2]$ is generated by $\phi_i(A,B)$, $i=1,2,3,4,5$.

Theorem 2.30. The functions $\phi_i(A,B)$, $i=1,2,3,4,5$ generate $(M_2 \times M_2)^G$ and are algebraically independent. Moreover these functions generate $[M_2 \times M_2]^G$.

Proof. Since the dimension of the generic orbit in $M_2 \times M_2$ is 3 the transcendence degree of $(M_2 \times M_2)^G$ is 5. As a generic orbit is determined by the values of $\phi_i(A,B)$, $1 \leq i \leq 5$ we easily deduce that $\phi_i(A,B)$, $i=1,\cdots,5$ are algebraically independent. Therefore these functions form a transcendental basis in $(M_2 \times M_2)^G$. Let $\phi \in (M_2 \times M_2)^G$. So ϕ is algebraic over $\phi_1,\cdots,\phi_5$. That is

$$\phi^m + \sum_{i=1}^{m} \rho_i(\phi_1,\cdots,\phi_5)\phi^{m-i} = 0. \tag{2.31}$$

Let $(A,B) \in U$. Then orb(A,B) is determined uniquely by $\phi_1,\cdots,\phi_5$. Combine this with (2.31) to deduce as in the proof of the previous theorem that $m=1$, i.e. ϕ is rational in $\phi_1,\cdots,\phi_5$. This shows that $\phi_1,\cdots,\phi_5$ generates $(M_2 \times M_2)^G$. Let $\phi \in [M_2 \times M_2]^G$. So ϕ is rational in $\phi_1,\cdots,\phi_5$. Also ϕ has a finite value for any $\phi_1,\cdots,\phi_5$. Thus ϕ must be a polynomial in $\phi_1,\cdots,\phi_5$. That is $[M_2 \times M_2]^G$ is generated by $\phi_1,\cdots,\phi_5$. ∎

We finally compare the results of this section to the classification procedure outlined in Section 1. We first note that if A and B are not commuting then the only non-trivial solution X to the system

$$AX-XA = 0, \quad BX-XB = 0 \tag{2.32}$$

is $X = \lambda I$. That is

$$M_2 \times M_2 - C = \mathcal{M}_{2,2,3}. \tag{2.33}$$

We conjecture

$$(2.34) \qquad M_2 \times M_2 - U = \mathcal{M}_{2,2,3,\rho_1}, \quad U - \mathcal{C} = \mathcal{M}_{2,2,3,\rho_2}, \quad \rho_1 > \rho_2.$$

It seems that for $(A,B) \in M_2 \times M_2 - U$ all 4×4 minors of $L(X,Y,A,B)$ are linearly independent. That is equivalent to the equality

$$\rho_1 = \binom{8}{4} = 70.$$

On $\mathcal{C} - W$ the system (2.32) has two independent solutions $X=A$, $X=B$ and $X=I$ if A or B are zero matrices. So

$$(2.35) \qquad \mathcal{C} - W = \mathcal{M}_{2,2,2}.$$

We conjecture

$$(2.36) \qquad V - W = \mathcal{M}_{2,2,2,\rho_3}, (U_1 \cup U_2) \cap \mathcal{C} - W = \mathcal{M}_{2,2,2,\rho_4}, \quad \rho_3 > \rho_4.$$

Finally

$$(2.37) \qquad W = \mathcal{M}_{2,2,0}.$$

We now verify Conjecture 1.26 for $m=n=2$.

<u>Theorem</u> 2.38. Let $A,B, \in M_2$ and assume $r = 4-\nu(A,B)$ where $\nu(A,B)$ is defined as in Section 1. Let $\mathcal{X}$ be the algebraic variety given by (1.9) ($m=n=2$). Then

$$\mathcal{X} = \overline{\mathrm{orb}}(A,B) \quad \text{for} \quad (A,B) \in M_2 \times M_2 - U$$

$$\mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2, \quad \mathcal{X}_1 = \overline{\mathrm{orb}}(A,B), \quad \dim \mathcal{X}_2 = 5, \quad \text{for } (A,B) \subset U -$$

$$(2.39) \qquad \mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2 \cup \mathcal{X}_3, \; \mathcal{X}_1 = \mathrm{orb}(A,B), \; \mathcal{X}_{i+1} = \{(\lambda_i I, \mu_i I)\}, \quad i = 1,2$$

$$\text{for} \quad (A,B) \in \mathcal{C} - V$$

$$\mathcal{X} = \overline{\mathrm{orb}}(A,B) \quad \text{for} \quad (A,B) \in V - W,$$

$$\mathcal{X} = \overline{\mathrm{orb}}(A,B) = \{(A,B)\}, \quad \text{for} \quad (A,B) \in W.$$

In particular Conjecture 1.26 is valid.

<u>Proof.</u> Assume first that $(A,B) \in M_2 \times M_2 - U$. It is enough to consider the case where A has two distinct eigenvalues. Thus we may assume that $A=D$, $B=E$ where D and E are of the form (2.4) and $e_{12}e_{21} \neq 0$. In particular

$\operatorname{rank} L(A,B,A,B) = 3$. Suppose that

(2.40) $$PX-XD = 0, \quad QX-XE = 0, \quad P = (p_{ij})_1^2, \quad Q = (q_{ij})_1^2$$

has a non-trivial solution X such that $|X| = 0$. By considering the matrices

$$P_1 = TPT^{-1}, \quad Q_1 = TQT^{-1}, \quad X_1 = TX$$

we may assume that either

(2.41) $$X = \begin{pmatrix} 1 & x \\ 0 & 0 \end{pmatrix}$$

or

(2.42) $$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Suppose that possibility (2.41) holds. Then the first equation of (2.40) yields

(2.43) $$p_{11} = \lambda_1, \quad p_{21} = 0, \quad x = 0.$$

The second equation of (2.40) implies that $e_{12} = 0$ which is impossible. So, if $X \neq 0$ satisfies (2.40) then $|X| \neq 0$. Hence (P,Q) is simultaneous similar to (D,E). That is $\mathcal{X} = \operatorname{orb}(A,B)$ and this orbit is closed.

Suppose next that (2.42) holds. Then the first equation of (2.40) implies

(2.44) $$p_{11} = \lambda_2, \quad p_{21} = 0.$$

The second equation of (2.40) yields $e_{21} = 0$ which is impossible. This proves the theorem for $(A,B) \in M_2 \times M_2 - U$. Assume now that $(A,B) \in U - C$. Again we may assume that A has two distinct eigenvalues. So let $A=D$, $B=E$, $e_{12} \neq 0$, $e_{21} = 0$.

Suppose that X is of the form (2.41). Then (2.43) holds. Then the second equality of (2.40) yields that $e_{12} = 0$ which is impossible. Assume now that X is of the form (2.42). So the equality (2.44) holds. Then the second equality of (2.40) implies

(2.45) $$q_{11} = e_{22}, \quad q_{21} = 0.$$

We claim (P,Q) is not similar to (D,E). Indeed if P,Q do not commute then P and Q have one common eigenvector corresponding to the eigenvalues λ_2 and e_{22} respectively. Since the common eigenvector of D and E corresponds to λ_1 and e_{11} respectively, (P,Q) and (D,E) are not similar. Clearly, if P and Q commute then (P,Q) and (D,E) are not similar. The set of all matrices (P,Q) satisfying (2.44) and (2.45) forms a manifold of dimension 4. Consider all pairs (P_1,Q_1) simultaneous similar to (P,Q). As the set of matrices (P,Q) satisfying (2.44) and (2.45) is invariant under the transformation (TPT^{-1}, TQT^{-1}), T is an upper triangular matrix. We deduce that the manifold X_2^0 consisting of matrices (P_1,Q_1) is of dimension 5. Thus X_2 = closure X_2^0 is an irreducible variety of dimension 5. This establishes the theorem in this case. Note that in this case

$$(2.46) \qquad X_1 \cap X_2 = \overline{\text{orb}}(D,E_1), \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad E_1 = \begin{pmatrix} e_{11} & 0 \\ 0 & e_{22} \end{pmatrix}.$$

Assume next that $(A,B) \in C - V$. So A and B are simultaneously similar to diagonal matrices such that at least one of them has distinct eigenvalues. Again we may assume $A=D$ has two distinct eigenvalues and $B=E$ with $e_{12} = e_{21} = 0$. Let X satisfy (2.40). Then as before either (2.41) and (2.43) holds or (2.42) and (2.44) holds. As $\nu(A,B) = 2$ we have an additional linearly independent matrix Y satisfying

$$(2.40)' \qquad PY-YD = 0, \quad QY-YE = 0, \quad Y = (y_{ij})_1^2.$$

We also assume that $|aX+bY| = 0$ for all a and b. By interchanging the roles of X and Y we get 3 possibilities

$$(2.47) \qquad X_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Y_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_2 = X, \quad Y_2 = Y_1^t,$$

$$X_3 = Y_1, \qquad Y_3 = I - X_1.$$

The choice X_1 and Y_1 is impossible since we get $\lambda_1 = \lambda_2$. The choice X_{i+1}, Y_{i+1} yields the solutions $P = \lambda_i I$ and $Q = e_{ii} I$ for $i=1,2$. So the theorem is established in this case.

Next let $(A,B) \in V-W$. By considering the matrices

$$A_1 = \alpha A+\beta B, \quad B_1 = \gamma A+\delta B, \quad \alpha\delta-\beta\gamma \neq 0$$

and making the similarity transformation (TAT^{-1}, TBT^{-1}) we may assume that

(2.48) $$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad B = \begin{pmatrix} \mu & 0 \\ 0 & \mu \end{pmatrix}.$$

Suppose that X is of the form (2.41). It is easy to check that it is impossible to satisfy the first equation of (2.40). So X must be of the form (2.42). Then the linearly independent matrix Y which satisfies (2.40)' can be chose of the form Y_3 given by (2.47).

So the equalities (2.40)-(2.40)' yield $P = \lambda I$, $Q=B$. Since the matrices

$$(A(\alpha) = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}, B) \in \mathrm{orb}(A,B)$$

for any $\alpha \neq 0$ we deduce that $(P,Q) \in \overline{\mathrm{orb}}(A,B)$. So $\overline{\mathrm{orb}}(A,B)$ is the only irreducible component of the system (1.9).

Assume finally that $(A,B) \in W$, i.e. $A = \lambda I$, $B = \mu I$. Then $\nu(\lambda I,\mu I) = 4$ so $r(\lambda I,\mu I) = 0$. But then $L(X_0,X_1,\lambda I,\mu I) = 0$ if and only if $X_0 = \lambda I$ and $X_1 = \mu I$. Thus we proved the equality (2.39). Suppose (A,B), (A_1,B_1) $\in \mathcal{M}_{2,2,r,\rho}$ and assume that the equality (1.24) holds. Then the variety $\mathcal{X}$ given by (1.9) contains the irreducible varieties $\overline{\mathrm{orb}}(A,B)$ and $\overline{\mathrm{orb}}(A_1,B_1)$ of dimension r. The equality (2.39) yield that $\mathcal{X}$ contains exactly one irreducible variety of dimension r. So $\overline{\mathrm{orb}}(A,B) = \overline{\mathrm{orb}}(A_1,B_1)$. Now Theorem 1.14 (the equality (1.16) yield that (A,B) and (A_1,B_1) are simultaneously similar. ∎

Remark 2.49. The proof of Theorem 2.38 yields the existence of non-similar pairs of matrices $(A_1,B_1),(A_2,B_2)$ such that

(2.50) $$\nu(A_1,B_1) = \nu(A_2,B_2) = \nu(A_1,B_1,A_2,B_2) = 1, \quad (A_1,B_1)\not\approx(A_2,B_2) \in U - \mathcal{C}.$$

This cannot happen in case that $m=0$ since the equality sign in (1.27) implies the similarity of A_0 and B_0 for any dimension n. (See Friedland [1980]).

3. Polynomial maps and varieties.

Let

$$\theta\ \mathbf{C}^\mu \to \mathbf{C}^\nu,\ \theta = (\theta_1(\zeta),\cdots,\theta_\nu(\zeta)), \quad \zeta \in \mathbf{C}^\mu$$

be a polynomial map, i.e. $\theta_i(\zeta)$ is a polynomial for $i = 1,\cdots,\nu$. In what follows we shall survey various properties of these maps needed in this paper. See van der Waerden [1950], Shafarevich [1974] and Whitney [1972) for general references on the algebraic and analytic properties of the polynomial maps. Specific results will be given the exact reference. The inverse image of ω, i.e. $\theta^{-1}(\omega)$ is called an algebraic variety in $\mathbf{C}^\mu$. That is any algebraic variety in $\mathbf{C}^\mu$ is given by some system of polynomial equations $\theta_i(\zeta) = \omega_i$, $i = 1,\cdots,\nu$, for some ν. We denote this algebraic variety by X. In what follows all the varieties mentioned here are algebraic. X is called reducible if $X = X_1 \cup X_2$ where $X_i \neq X$ for $i=1,2$ and each X_i is a variety. Otherwise X called irreducible. It is well known that any variety is a finite union of irreducible ones. Any irreducible variety is connected. A point x in an irreducible variety X is called regular if in the neighborhood of this point, X is a manifold. The dimension of this manifold does not depend on a choice of the regular point and it is called the dimension of X. Let X^0 be the set of all regular points in X. Then X^0 is an open connected set in X. In particular X^0 is a manifold. Recall that Y is an open set in X if $Y = X \cap W$ for some open set W in $\mathbf{C}^\mu$. Y is called algebraically open if W is the set of points for which $p(\zeta) \neq 0$, for some polynomial p. Note that a finite union of open algebraic sets $\bigcup_{\alpha=1}^{m} W_\alpha$ (defined by $p_\alpha \neq 0$) is an open algebraic set W given by $p = p_1\cdots p_m$. If Y is an open algebraic set in irreducible variety X then Y is connected. Let X be an irreducible and X^0 be the manifold of the regular points of X.

Then X^0 is an open algebraic set. That is the set of singular points in X is a subvariety of X. For a reducible variety X the dimension of X is defined as the largest dimension of its irreducible components. A variety X is called homogeneous if $tX = X$ for all $t \neq 0$. In that case X is the zero set of $\theta_i(\zeta) = 0$, $i=1,\cdots,\nu$, where each $\theta_i(\zeta)$ is a homogeneous polynomial. For any variety X, let X^h be the homogeneous variety given as the closure of the union of tX, for all possible t. Clearly X is irreducible if and only if X^h is irreducible. To study the homogeneous varieties it is convenient to introduce the projective space $\mathbf{P}^{\nu-1}$. It is the set $\mathbf{C}^\nu - \{0\}$ when one identify ζ with $t\zeta$ for any $t \neq 0$. Thus any homogeneous variety X of dimension d gives rise to the projective variety $\hat{X}$ of dimension $d-1$ in $\mathbf{P}^{\nu-1}$ and vice versa. A well known result for projective varieties claims (e.g. Shafarevich [1974]).

<u>Theorem</u> 3.1. Let $\hat{X}_1$ and $\hat{X}_2$ be projective varieties of dimensions d_1-1 and d_2-1 respectively in $\mathbf{P}^{\nu-1}$. If $d_1+d_2 \geq \nu+1$ then $\hat{X}_1 \cap \hat{X}_2$ is a non-empty projective variety at least of dimension $\nu-d_1-d_2+1$.

Let $\hat{X}$ be a projective variety of dimension $d-1$. Let $\hat{Y}$ be a projective variety given by the intersection of $d-1$ hyperplanes

$$(3.2) \qquad H_i = \{x, \sum_{j=1}^{\nu} a_{ij}x_j = 0, \quad x \in \mathbf{P}^{\nu-1}\}, \quad i=1,\cdots,d-1.$$

Then, according to Theorem 3.1 $\hat{X} \cap \hat{Y}$ is a non-empty projective variety. We assume here that $(a_{i1},\cdots,a_{i\nu}) \in \mathbf{P}^{\nu-1}$, $i=1,\cdots,d-1$. Moreover except for some variety Z in

$$\mathbf{P}^{\nu-1,d-1} = \underbrace{\mathbf{P}^{\nu-1} \times \cdots \times \mathbf{P}^{\nu-1}}_{d-1}$$

$\hat{X} \cap \hat{Y}$ consists of exactly δ points in $\mathbf{P}^{\nu-1}$. The number δ is called the degree of the variety $\hat{X}$ and is denoted by $\deg \hat{X}$. In what follows we need the following result.

Theorem 3.3. Let $\hat{X}$ be an irreducible projective variety in P^{d-1} of dimension $d-1 \geq 1$ and degree δ. Let $H_1,\cdots,H_{d-1}$ be any $d-1$ hyperplanes in $P^{\nu-1}$. Let

$$\hat{X}_k = \hat{X} \bigcap_{j=1}^{k} H_j = \bigcup_{i=1}^{\nu_k} \hat{X}_{ki}, \quad 1 \leq k \leq d-1, \tag{3.4}$$

$\hat{X}_{k1},\cdots,\hat{X}_{k\nu_k}$ are the irreducible components of $\hat{X}_k$. Then

$$\dim(\hat{X}_k) \geq d-1-k \tag{3.5}$$

and

$$\sum_{i=1}^{\nu_k} \deg \hat{X}_{ki} \leq \deg \hat{X} = \delta. \tag{3.6}$$

Proof. Consider

$$\hat{X}_1 = \hat{X} \cap H_1 .$$

If $\hat{X} \subseteq H_1$ then $\hat{X}_1 = \hat{X}$, $\nu_1 = 1$ and the theorem trivially holds. If $\hat{X}$ does not lie entirely in H_1 then

$$\hat{X}_1 = \bigcup_{i=1}^{\nu_1} \hat{X}_{1i}$$

and

$$\dim \hat{X}_{1i} = d-2, \quad i = 1,\cdots,\nu_1.$$

So

$$\hat{X}_1 \bigcap_{j=2}^{d-1} H_j = \cup(\hat{X}_{i1} \bigcap_{j=2}^{d-1} H_j)$$

and for generic $H_2,\cdots,H_{d-1}$, $\hat{X}_{1i} \bigcap_{j=2}^{d-1} H_j$ will consist of $\deg\hat{X}_{1i}$ distinct points.

Since $\hat{X}_{1i} \neq \hat{X}_{i\ell}$ for $i \neq \ell$ for the generic hyperplanes $H_2,\cdots,H_{d-1}$ the sets $\hat{X}_{1i} \bigcap_{j=2}^{d-1} H_j$ and $\hat{X}_{1\ell} \bigcap_{j=2}^{d-1} H_j$ must be distinct. Otherwise $\hat{X}_{1i} \cap \hat{X}_{1\ell}$ will contain a smooth manifold of dimension $d-2$ and thus $\hat{X}_{1i} = \hat{X}_{1\ell}$, as X_{1i} and $\hat{X}_{1\ell}$ are irreducible which is impossible. Hence in case $\hat{X} \not\subseteq H_1$ we have

$$\deg \hat{X} = \sum_{i=1}^{\nu_1} \deg\hat{X}_{1i}.$$

Consider now $\hat{X} \cap H_1 \cap H_2$. Then

$$\hat{X} \cap H_1 \cap H_2 = \left(\bigcup_{i=1}^{\nu_1} \hat{X}_{1i}\right) \cap H_2.$$

Also

$$\hat{X}_{1i} \cap H_2 = \bigcup_{j=1}^{\mu_i} \hat{X}_{1ij}$$

As $\dim \hat{X}_{1i} \geq d-2$ we have $\dim \hat{X}_{1ij} \geq d-3$. According to what we proved above

$$\deg \hat{X}_{1i} = \sum_{j=1}^{\mu_i} \deg X_{1ij}.$$

So

$$\deg \hat{X} = \sum_{i=1}^{\nu_1} \deg \hat{X}_{1i} = \sum_{i=1}^{\nu_1} \sum_{j=1}^{\mu_i} \deg X_{1ij}. \tag{3.7}$$

Also

$$\hat{X}_2 = \bigcup_{i=1}^{\nu_1} \bigcup_{j=1}^{\mu_i} \hat{X}_{1ij}.$$

Consider the decomposition (3.4) of $\hat{X}_2$ to irreducible factors. Clearly $\hat{X}_{2\alpha} = \hat{X}_{1ij}$ for some i,j. However if $\hat{X}_{1pq} \subseteq \hat{X}_{1ij}$ then $\hat{X}_{1pq}$ will not appear in the irreducible decomposition of $\hat{X}_2$. Thus the equality (3.7) implies the inequality (3.6). This completes the proof of the theorem for $k=2$. The same arguments establishes the theorem for any $3 \leq k \leq d-1$. ∎

<u>Remark</u> 3.8. It is easy to show that for generic hyperplanes $H_1,\cdots,H_k$ the equality sign would hold in (3.5) and (3.6). Also, the degrees of all $\hat{X}_{ki}$ must be the same.

Let $M_{n,m}$ be the set of $n \times m$ complex valued matrices. We identify $M_{n,m}$ with $\mathbb{C}^{mn}$. For simplicity of notations we shall assume

$$1 \leq m \leq n. \tag{3.9}$$

For $1 \leq r \leq m \leq n$ let $M_{n,m,r}$ be the set of matrices A such that rank $A \leq r$. That is

$$M_{n,m,r} = \{A,\ A[\alpha|\beta] = 0,\quad \alpha \in Q_{r+1,n},\ \beta \in Q_{r+1,m}\}. \tag{3.10}$$

Hence $M_{n,m,r}$ is a homogeneous variety. Thus we can view $\hat{M}_{n,m,r}$ as a projective variety in P^{mn-1}. We claim that $M_{n,m,r}$ is an irreducible variety. Let $M^0_{n,m,r}$ be a set of all $A \in M_{n,m}$ with $\operatorname{rank} A = r$. Clearly this set is open in $M_{n,m,r}$. Also for $A \in M^0_{n,m,r}$

$$(3.11) \qquad A = PD_rQ, \; P \in GL_n, \quad Q \in GL_m,$$

where $D_r = (d_{ij})$, $d_{ii} = 1$, $i = 1,\cdots,r$ and all other entries of D_r vanish. As GL_n and GL_m are connected manifolds $M^0_{n,m,r}$ is a connected manifold. So $M_{n,m,r}$ is irreducible. We claim

$$(3.12) \qquad \dim M_{n,m,r} = r(n+m-r) \quad (r \leq m \leq n).$$

Indeed, pick up $A_0 \in M^0_{n,m,r}$. Assume for simplicity that the first r rows of A_0 are linearly independent. Thus, if the first r rows of A stay in the neighborhood of the first rows of A_0 these rows will be linearly independent. This gives rm independent parameters. Now, if $A \in M_{n,m,r}$ then any other row of A is an arbitrary linear combination of the first r rows of A. This gives us additional $r(n-r)$ parameters.

So we proved (3.12). Next we claim

<u>Theorem</u> 3.13. Let $M_{n,m,r}$ be the irreducible variety of all $n\times m$ matrices with the rank r at most. Suppose that $1 \leq r < m \leq n$. Then

$$(3.14) \qquad \deg \hat{M}_{n,m,r} = (r+1)^{(m-r)(n-r)}.$$

To prove this theorem as well as other results we need the following. Let $\theta : \mathbf{C}^\mu \to \mathbf{C}^\nu$ be a polynomial map. Denote by $\partial\theta(\zeta)$ the Jacobian of θ at the point ζ. Let ρ be the rank of $\partial\theta$. A point $\zeta \in \mathbf{C}^\mu$ is called regular if $\operatorname{rank}\partial\theta(\zeta) = \rho$. A point ζ is called singular or critical if $\operatorname{rank}\partial\theta(\zeta) < \rho$. Let S be the set of all singular points of θ. Clearly S is an algebraic variety. So $\mathbf{C}^\mu - S$ the set of regular points is an open algebraic set in $\mathbf{C}^\mu$. Therefore $\theta(\mathbf{C}^\mu - S)$ is a connected manifold in $\mathbf{C}^\nu$ of dimension ρ. Hence, $X = c\ell\,\theta(\mathbf{C}^\mu)$ is an irreducible variety of dimension ρ in $\mathbf{C}^\nu$. Also

$\theta(\mathbf{C}^{\mu} - S)$ is an open algebraic set in X. So $X - \theta(\mathbf{C}^{\mu} - S)$ is a closed subvariety in X. Thus, there exists a non-trivial polynomial κ such that

$$\kappa(\omega) = 0 \tag{3.15}$$

for any $\omega \in X - \theta(\mathbf{C}^{\mu} - S)$. In particular any $\omega \in X - \theta(\mathbf{C}^{\mu})$ (ω is an omitted value in X) must satisfy (3.15). Consider $Y = c\ell\ \theta(S)$. It follows that Y is a subvariety of X. So there exists another non-trivial polynomial κ such that any critical value ω of θ must satisfy (3.15). Let ω be a noncritical value, i.e. $\omega \in \theta(\mathbf{C}^{\mu} - S)$. Then

$$\theta^{-1}(\omega) = \bigcup_{j=1}^{m} Y_j \tag{3.16}$$

where each Y_j is an irreducible variety of the dimension $\mu - \operatorname{rank}\partial\theta$. The number m is independent of the point ω and is called the degree of the map - deg θ. The continuity argument implies that whenever we have the decomposition (3.16) to irreducible varieties then $m \leq \deg\theta$. The map θ is called regular if

$$\operatorname{rank} \partial\theta = \min(\mu, \nu).$$

The arguments above show

<u>Theorem</u> 3.17. Let $\theta : \mathbf{C}^{\mu} \to \mathbf{C}^{\nu}$ be a polynomial map. Let the degree of θ - $\deg\theta$ - be defined as above. Consider decomposition (3.16) of $\theta^{-1}(\omega)$ to its irreducible varieties. Then $m \leq \deg\theta$. If $\omega \in \theta(\mathbf{C}^{\mu})$ is not a critical value then $m = \deg\theta$ and $\dim Y_i = \mu - \operatorname{rank}\partial\theta$, $i = 1,\cdots,\deg\theta$. Assume that $\mu \geq \nu$ and suppose that θ is a regular map. Then there exists a non-trivial polynomial κ such that any omitted value $\omega \notin \mathbf{C}^{\nu} - \theta(C^{\mu})$ must satisfy (3.15). Assume furthermore that $\mu = \nu$. Then any non-critical value $\omega \in \mathbf{C}^{\nu}$ is obtained exactly $\deg\theta$ times.

Note that if θ is onto map then $\kappa(\omega)$ can be chosen to be $\equiv 1$. The most interesting case is when $\mu = \nu$. In that case there are simple sufficient conditions for θ to be an onto map. Let $p(\zeta)$ be a polynomial in μ variables. Denoted be $\deg p$ the degree of p. Define

$$(3.18) \qquad p_\pi(\zeta) = \lim_{t\to\infty} \frac{p(t\zeta)}{t^{\deg p}}, \ \deg p \geq 1, \qquad p_\pi(\zeta) = 0 \quad \text{if } p \text{ is a constant.}$$

Put

$$(3.19) \qquad \theta_\pi = (\theta_{1\pi}, \cdots, \theta_{\nu\pi}), \quad \deg\theta = \prod_{i=1}^{\nu} \deg \theta_i.$$

Theorem 3.20. Let $\theta : \mathbf{C}^\mu \to \mathbf{C}^\mu$ be a polynomial map. Assume that the system

$$(3.21) \qquad \theta_\pi(\zeta) = 0$$

has the only solution $\zeta = 0$. Then the system

$$(3.22) \qquad \theta(\zeta) = \omega$$

is always solvable. The number of distinct solutions of (3.22) is at most $\deg\theta$. Moreover, there exists a non-trivial polynomial $\kappa(w)$, $w \in \mathbf{C}^\mu$ such that the equation (3.22) has exactly $\deg\theta$ distinct solutions unless ω satisfies (3.15). That is θ is $\deg\theta$ covering of $\mathbf{C}^\mu$.

Remark 3.23. This theorem is essentially due to Noether and van der Waerden [1928]. It was rediscovered by us in Friedland [1977] (Theorem 2.1). Although Theorem 2.1 is stated in a slightly different form we did show in the proof of Theorem 2.1 that the equation (3.21) has exactly $\deg\theta$ distinct solutions unless ω is a critical value. In that case the arguments preceeding Theorem 3.17 imply that ω must satisfy a non-trivial equation (3.15). In fact $\kappa(\omega)$ is non-constant if $\deg\theta > 1$.

Proof of Theorem 3.13. Consider generic $H_1, \cdots, H_{p-1}$ hyperplanes where $p = r(n+m-r)-1$ in the projective space $\mathbf{P}^{mn-1}$. So

$$\hat{M}_{n,m,r} \bigcap_{j=1}^{p} H_j = \{A_1,\cdots,A_q\}$$

where $q = \deg(\hat{M}_{n,m,r})$. When we vary $(H_1,\cdots,H_p) \in \underbrace{\mathbf{P}^{m\nu-1} \times \cdots \times \mathbf{P}^{m\nu-1}}_{p}$ the points $\{A_1,\cdots,A_q\}$ sweep an open set in $\hat{M}_{n,m,r}$. So we may assume that $H_1,\cdots,H_p$ were chosen such that

(i) $A_1,\cdots,A_q$ are pairwise distinct

(ii) the $r\times r$ minors composed of the first rows and columns of $A_1,\cdots,A_q$ are distinct from zero. Also (1,1) entry of each A_i is different from zero. As each A_i lies in the projective space we may assume that (1,1) of each A_i is equal to 1.

Then $A_1,\cdots,A_q$ are the solutions of the following mn polynomial equations whose variables are the entries of $X = (x_{ij})$, $i=1,\cdots,n$, $j=1,\cdots,m$.

(I) all $(r+1)\times(r+1)$ minors of X which include the first r rows and columns of X are equal to zero.

(II) $H_j(X) = 0$, $j = 1,\cdots,P$

(III) $x_{11} = 1$.

The set of equations (I) consists of $(m-r)(n-r)$ homogeneous equations of degree $r+1$. Sets II + III consist of $r(n+m-r)$ linear equations. So we have exactly mn equations. The left hand side of I, II and III defines a map $\theta : \mathbf{C}^{mn} \to \mathbf{C}^{mn}$ of the degree $\delta = (r+1)^{(m-r)(n-r)}$. Next we show that the system (3.21) has the only solution $X=0$. Otherwise at least one A_i will have (1,1) zero entry which contradicts our assumption.

Theorem 3.20. yields that $q \leq \delta$. It is left to show that we have the equality sign $q = \delta$. As $\hat{M}_{n,m,r}$ is an irreducible variety such that $\hat{M}^0_{n,m,r}$ is open in $\hat{M}_{n,m,r}$ it follows that p generic hyperplanes $H_1,\cdots,H_p$ will intersect $\hat{M}^0_{n,m,r}$ transversally. That is we may assume that θ is regular at the points $A_1,\cdots,A_q$. In that case the Remark 3.23 implies that $q=\delta$ and the theorem is proved. ∎

In $M_{(m+1)n^2,n^2}$ consider a subspace L of matrices of the form $L(X_0,\cdots,X_m,\lambda A_0,\cdots,\lambda A_m)$ where $X_0,\cdots,X_m$ are arbitrary $n\times n$ matrices and λ is a complex parameter. Let

$$\text{(3.24)} \qquad Y = M_{(m+1)n^2,n^2,r} \cap L \quad , \quad Y = \bigcup_{j=1}^{\eta} Y_j$$

where each Y_j is an irreducible homogeneous variety. According to Theorem 3.13

$$\deg \hat{M}_{(m+1)n^2,n^2,r} = (r+1)^{[(m+1)n^2-r][n^2-r]} = \kappa(n,m,r).$$

Next Theorem 3.3 yields

$$\text{(3.25)} \qquad \eta \leq \sum_{j=1}^{\eta} \deg \hat{Y}_j \leq \kappa(n,m,r).$$

Let X be the variety given by (1.9). Assume that (1.10) is its decomposition to irreducible varieties in $C^{(n+1)n^2}$. Clearly each irreducible X_i is obtained by restricting some Y_j to the hyperplane H of matrices of the form $L(X_0,\cdots,X_m,A_0,\cdots,A_m)$. Since some Y_j may have an empty intersection with H we have the inequality $\kappa \leq \eta$.

This establishes (1.11). More precisely, Theorem 3.3 implies

$$\text{(3.26)} \qquad \sum_{j=1}^{\kappa} \deg \hat{X}_j^h \leq \kappa(n,m,r).$$

Thus if we can compute the degree of $\widehat{\mathrm{orb}}(A_0,\cdots,A_m)^h$ we could probably improve the inequality (1.11) by means of the above inequality.

4. Algebraic functions.

Let $p(\lambda,x)$ be a polynomial of the form

(4.1) $$p(\lambda,x) = \lambda^n + \sum_{i=1}^{n} p_i(x)\lambda^{n-i}.$$

As usual let $C[\lambda,x]$ be the ring of polynomials in λ and x. As $C[\lambda,x]$ is UFD (unique factorization domain) $p(\lambda,x)$ decomposes to irreducible factors

(4.2) $$p(\lambda,x) = \prod_{i=1}^{k} q_i(\lambda,x)^{\mu_i},$$

where each $q_i(\lambda,x)$ is irreducible and q_i and q_j are coprime for $i \neq j$. Moreover since $p(\lambda,x)$ is monic in λ each $q_i(\lambda,x)$ can be assumed to be monic in λ of degree 1 at least in λ variable. For this property and others consult for example with Whitney [1972]. We call $p(\lambda,x)$ degenerated if in the decomposition (4.2) some factor is repeated twice at least. That is $\mu_i > 1$ for some i. In what follows we consider the roots of $p(\lambda,x)$

(4.3) $$p(\lambda,x) = 0.$$

Then each root $\lambda(x)$ is an algebraic function of x. We can name these roots by $\lambda_1(x),\cdots,\lambda_n(x)$. Clearly there are two identical roots $\lambda_i(x)$ and $\lambda_j(x)$ if and only if $p(\lambda,x)$ is degenerated. For $\lambda_i(x)$ a point ζ is called a regular point (point of analyticity) if $\lambda_i(x)$ is analytic in the neighborhood of this point. Otherwise ζ is called a branch (singular) point. Let ζ be a branch point. Then it is possible to group the eigenvalues as follows

(4.4) $$\{\lambda_1(x),\cdots,\lambda_{m_1}(x)\}, \{\lambda_{m_1+1}(x),\cdots,\lambda_{m_2+1}(x),\cdots\{\lambda_{m_{\kappa-1}+n}(x),\cdots,\lambda_{m_\kappa}(x)\}$$

such that, when we circle once on the small circle around ζ each group of eigenvalues undergoes a cyclic permutation. For brevity each group will be called a cycle at ζ and the number of elements of a cycle will be called its period. So if the period of a given cycle is one then the corresponding $\lambda_i(x)$ is analytic in the neighborhood of ζ. Thus ζ is a branch point if

there exist a cycle of eigenvalues with the period greater than one. Assume that $\{\lambda_1(x),\cdots,\lambda_m(x)\}$ form a cycle. Then we have the Puiseaux series

$$(4.5) \qquad \lambda_h(x) = \sum_{j=0}^{\infty} \mu_j \omega^{j(h-1)}(x-\zeta)^{j/m}, \quad h = 1,\cdots,m,$$

where $\omega = e^{2\pi i/m}$. See for example Whitney [1972, p. 32].

It is well known that the equation $p(\lambda,x) = 0$ has a finite number of branch points $\zeta_1,\cdots,\zeta_b$. So each $\lambda_i(x)$ is a multivalued analytic function on

$$(4.6) \qquad D = C - \{\zeta_1,\cdots,\zeta_b\}.$$

Suppose that $p(\lambda,x)$ is irreducible. Then each $\lambda_i(x)$ is n valued. In other words each $\lambda_i(x)$ is analytic univalued on n cover sheets of D. More precisely, starting from one branch of $\lambda_i(x)$ it is possible to recover all other branches of $\lambda(x)$ on D by analytic continuation. Using the decomposition (4.2) we can find how many branches each $\lambda_i(x)$. has. In particular, if all the eigenvalues $\lambda_i(x)$, $i=2,\cdots,n$ can be generated from $\lambda_1(x)$ then $p(\lambda,x)$ must be irreducible. Next we study what happens when $x \to \infty$. Assume that $p(\lambda,x)$ is of the form (4.1). Define

$$(4.7) \qquad \delta = \max_{1\leq i\leq n} \deg p_i/i.$$

Divide the equation (4.1) by $x^{n\delta}$ to deduce that

$$(4.8) \qquad |\lambda_i(x)| \leq K|x|^{\delta} \quad \text{for} \quad |x| \geq r.$$

So at $x = \infty$, i.e. $|x| > r$ for some large r, we divide the eigenvalues to the cycles of the form (4.4). Now the group $\lambda_1,\cdots,\lambda_m$ must have the expansion

$$(4.9) \qquad \lambda_h(x) = \omega^{\alpha h} x^{\alpha/m}\left(\sum_{j=0}^{\infty} \nu_j\, \omega^{jh} x^{-j/m}\right), \quad 0 \leq \alpha \leq m\delta.$$

The case $\alpha=0$ can only correspond to the case where $m=1$ and $\lambda_1(x)$ is a constant function.

Lemma 4.10. Let $q(\lambda,x)$ be of the form (4.1)

$$q(\lambda,x) = \lambda^m + \sum_{i=1}^{m} q_i(x)\lambda^{m-i}. \tag{4.11}$$

Assume that $q(\lambda,x)$ divides $p(\lambda,x)$. Then

$$\deg q_i(x) \leq i\delta, \quad i=1,\cdots,m. \tag{4.12}$$

Proof. It is possible to rename the roots $\lambda_1,\cdots,\lambda_n$ of $p(\lambda,x) = 0$ such that

$$q(x) = \prod_{i=1}^{m} (\lambda-\lambda_i(x)). \tag{4.13}$$

According to (4.9) $|q_i(x)| \leq K|x|^{i\delta}$ for $|x|$ big enough. This proves (4.12). ■

Definition 4.14. A point ζ is called an intersection point if either ζ is a branch point or there exist two distinct eigenvalues $\lambda_i(x)$ and $\lambda_j(x)$ (in the neighborhood of ζ) analytic at ζ such that $\lambda_i(\zeta) = \lambda_j(\zeta)$ if ζ is finite. (In case that $\zeta = \infty$ we demand that $\lim_{x\to\infty} \frac{\lambda_i(x)}{\lambda_j(x)} = 1$.)

Consider for example

$$p(\lambda,x) = \prod_{i=1}^{n} [\lambda - (a_i + b_i x^r], \quad r \text{ is a positive integer.} \tag{4.15}$$

Then we do not have any branch points but there are intersection points

$$\zeta_{ij}^r = (a_i-a_j)/(b_j-b_i), \quad 1 \leq i < j \leq n \tag{4.16}$$

if $(a_i,b_i) \neq (a_j,b_j)$.

Suppose that $p(\lambda,x)$ is not degenerate. In order to find the intersection points we consider

$$D(p) = \prod_{1\leq i<j\leq n} (\lambda_i(x)-\lambda_j(x))^2 = D(x). \tag{4.17}$$

Since $p(\lambda,x)$ is monic in λ it is well known that $D(p)$ is a polynomial in $p_1,\cdots,p_n$. The expansion (4.9) implies that

$$(4.18) \qquad \deg D(x) \le n(n-1)\delta.$$

Clearly $D(x) \equiv 0$ if and only if $p(\lambda,x)$ is degenerate. Thus if $p(\lambda,x)$ is nondegenerate then a finite point ζ is an intersection if and only if $D(\zeta) = 0$. We now study the connection between the multiplicity of the root $x = \zeta$ in $D(x)$ and the nature of the intersection point ζ.

<u>Theorem</u> 4.19. Let ζ be a finite intersection point. Suppose that $\lambda_1(x), \cdots, \lambda_n(x)$ break up to the cycles as given by (4.4). Assume furthermore that

$$(4.20) \qquad \begin{aligned} &\lambda_1(\zeta) = \cdots = \lambda_{m_{k_1}}(\zeta) \ne \lambda_{m_{k_1+1}}(\zeta) = \cdots = \lambda_{m_{k_2}}(\zeta) \\ &\ne \lambda_{m_{k_{\theta-1}}+1}(\zeta) = \cdots = \lambda_{m_{k_\theta}}(\zeta), \qquad k_\theta = \kappa. \end{aligned}$$

Then $x=\zeta$ is a root of $D(x)$ at least of multiplicity

$$(4.21) \qquad L = \sum_{i=1}^{\kappa} (m_i - 1) + 2\sum_{p=0}^{\theta-1} \sum_{k_p+1 \le i<j \le k_{p+1}} \min(m_i, m_j), \quad k_0 = 0,\ k_\theta = \kappa.$$

<u>Proof.</u> Assume that $\lambda_1(x), \cdots, \lambda_m(x)$ is a cycle of period m at $x=\zeta$. Consider the product

$$f(x) = \prod_{1 \le i<j \le m} (\lambda_i(x) - \lambda_j(x))^2.$$

By completing one circle on $|z-\zeta| = r$ we permute $\lambda_1, \cdots, \lambda_m$ cyclically so $f(x)$ is analytic and univalued in $|z-\zeta| < r$. Each $\lambda_k(x)$ has an expansion (4.5). So $\lambda_i(x) - \lambda_j(x)$ is divided by $(x-\zeta)^{1/m}$. Thus $f(x)$ is divided by $(x-\zeta)^{m-1}$. This shows that $D(x)$ is divided by $(x-\zeta)^\mu$, $\mu = \sum_{i=1}^{\kappa} (m_i - 1)$.

Assume now that $\lambda_{m+1}(x), \cdots, \lambda_{m+\ell}(x)$ is another cycle of period ℓ such that $\lambda_i(\zeta) = \lambda_j(\zeta)$, $i,j = 1, \cdots, m+\ell$.

Consider the function

$$g(x) = \prod_{1 \le i \le m,\, m+1 \le j \le m+\ell} (\lambda_i(x) - \lambda_j(x)).$$

By circling around the point ζ we permute the branches $\lambda_1, \cdots, \lambda_m$ and $\lambda_{m+1}, \cdots, \lambda_{m+\ell}$. So $g(x)$ remains univalued in the neighborhood of ζ. That is $g(x)$ is analytic for $|x-\zeta| < r$. Without loss in generality we may assume that $m \ge \ell$. Then $g(x)$ is divided by $[(x-\zeta)^{1/m}]^{m\ell} = (x-\zeta)^\ell$.

Thus $g^2(x)$ is divided by $(x-\zeta)^{2\ell}$. So altogether $D(x)$ is divided by $(x-\zeta)^L$ where L is given by (4.21). ■

Theorem 4.22. Let ζ be a finite intersection point. Assume furthermore that ζ is a simple root of $D(x)$. Then ζ is a branch point of the algebraic function $p(\lambda,x) = 0$ and the equation $p(\lambda,\zeta) = 0$ has exactly $n-1$ distinct roots. That is at $x=\zeta$ one has exactly one cycle $\{\lambda_1(x),\lambda_2(x)\}$ of period two while all other branches $\lambda_3(x),\cdots,\lambda_n(x)$ are analytic in the neighborhood of ζ. Moreover in the expansion (4.5) $\mu_1 \neq 0$ $(m=2)$.

Proof. Our assumptions imply that L given by (4.21) is equal to 1. So $k_i = i$, $i=0,\cdots,\kappa$ and $m_1=2$, $m_i=1$, $i=2,\cdots,\kappa$. Thus ζ is a branch point and $p(\lambda,\zeta) = 0$ has exactly $n-1$ distinct roots. Furthermore

$$(4.23) \qquad (\lambda_1(x)-\lambda_2(x))^2 = 4(x-\zeta)\Big(\sum_{j=0}^{\infty} \mu_{2j+1}(x-\zeta)^j\Big)^2.$$

Thus is $\mu_1=0$ then $x=\zeta$ is a zero of order 3 at least of $D(x)$. Therefore $\mu_i \neq 0$. ■

Theorem 4.22 shows that the simple zeros of $D(x)$ describe the simplest possible branch points of $p(\lambda,x) = 0$. Theorem 4.19 enables us to analyze the double roots of $D(x)$.

Theorem 4.24. Let ζ be a finite intersection point. Assume furthermore that ζ is a double root of $D(x)$. Then one of the following conditions holds.

(i) $p(\lambda,\zeta) = 0$ has $n-1$ distinct roots, i.e.

$$\lambda_1(\zeta) = \lambda_2(\zeta), \quad \lambda_i(\zeta) \neq \lambda_j(\zeta) \quad \text{for} \quad 2 \leq i < j \leq n.$$

In that case each $\lambda_i(x)$ is analytic in the neighborhood of ζ and $\lambda_1'(\zeta) \neq \lambda_2'(\zeta)$.

(ii) $p(\lambda,\zeta) = 0$ has $n-2$ distinct roots and one of them is triple, i.e.

$$\lambda_1 = (\zeta) = \lambda_2(\zeta) = \lambda_3(\zeta), \quad \lambda_i(\zeta) \neq \lambda_j(\zeta) \quad \text{for} \quad 3 \leq i < j \leq n.$$

In that case $x = \zeta$ is a branch point with exactly one cycle $\{\lambda_1(x), \lambda_2(x), \lambda_3(x)\}$, while all other branches $\lambda_4(x), \cdots, \lambda_n(x)$ are analytic. Moreover in the expansion (4.5) $\mu_1 \neq 0$ $(m=3)$.

(iii) $p(\lambda, \zeta)$ has $n-2$ distinct roots and two of them are double, i.e. $\lambda_1(\zeta) = \lambda_2(\zeta)$, $\lambda_{n-1}(\zeta) = \lambda_n(\zeta)$, $\lambda_i(\zeta) \neq \lambda_j(\zeta)$ for $2 \leq i < j \leq n-1$. In that case $x=\zeta$ is a branch point with exactly two cycles $\{\lambda_1(x), \lambda_2(x)\}$, $\{\lambda_{n-1}(x), \lambda_n(x)\}$ while all other branches $\lambda_3(x), \cdots, \lambda_{n-2}(x)$ are analytic. Moreover, in the expansions (4.5) for $\{\lambda_1(x), \lambda_2(x)\}$ and $\{\lambda_{n-1}(x), \lambda_n(x)\}$ $\mu_1 \neq 0$.

The proof of Theorem 4.24 is quite analogous to the proof of Theorem 4.22 and uses only Theorem 4.19. So we omit its proof.

5. Special polynomials in two variables.

In this section we consider polynomials $p(\lambda,x)$ of the form

$$(5.1)\quad p(\lambda,x) = \lambda^n + \sum_{i=1}^{n} p_i(x)\lambda^{n-i},\ p_i(x) = \sum_{j=0}^{ir} p_{ij}x^j,\ i=1,\cdots,n.$$

Here r is a positive integer. Such a polynomial $p(\lambda,x)$ is uniquely determined by the coefficient vector

$$(5.2)\qquad p = (p_{10},p_{11},\cdots,p_{1r},\cdots,p_{n0},\cdots,p_{n(nr)}).$$

We shall identify $p(\lambda,x)$ with its coefficient vector and no ambiquity will arise. Thus any polynomial $p(\lambda,x)$ is given by a point p in

$$P^{n,r} = C^{[(n+1)r+2]n/2}.$$

We next consider the algebraic function $\lambda(x)$ given by the equation $p(\lambda,x) = 0$. In that case δ defined by (4.7) is at most r. So each $\lambda(x)$ has the Puiseaux expansion

$$(5.3)\qquad \lambda(x) = x^r \sum_{j=0}^{\infty} \nu_j x^{-j/m}$$

around $\zeta = \infty$. Here ν_0 is the root of the equation

$$(5.4)\qquad \nu^n + \sum_{i=1}^{n} p_{i(ir)}\nu^{n-i} = 0.$$

Thus $\zeta = \infty$ is an <u>intersection</u> point if (5.4) has at least one double root. Let $D(x)$ be given by (4.17). According to (4.18) the degree of $D(x)$ is at most $n(n-1)r$.

So

$$(5.5)\qquad D(x) = \sum_{i=0}^{n(n-1)r} d_i(p)x^{rn(n-1)-i}$$

where each $d_i(p)$ is a polynomial in the coefficint vector p. In particular

$$(5.6)\qquad d_0(p) = \prod_{1\leq i<j\leq n} (\nu_i - \nu_j)^2$$

where $\nu_1,\cdots,\nu_n$ are the roots of (5.4). So $d_0(p)$ is a polynomial in $p_{11},\cdots,p_{n(nr)}$ which is called the discriminant of

$$(5.7)\qquad p_n(\lambda) = \lambda^n + \sum_{i=1}^{n} p_{i(ir)}\lambda^{n-i}.$$

Thus for all $p \in P^{n,r}$ such that $d_0(p) \neq 0$ the polynomial $D(x)$ is exactly of degree $rn(n-1)$. Let $p(\lambda,x)$be given by (4.15). Then

$$(5.8) \qquad D(x) = \prod_{1 \le i<j \le n} [(b_i - b_j)x^r + (a_i - a_j)]^2.$$

Suppose that $b_i \neq b_j$ and $a_i \neq a_j$ for $i \neq j$. Then $D(x)$ is a polynomial of degree $rn(n-1)$. Clearly we can choose (a_i, b_i), $i=1,\cdots,n$ such that we would have exactly $rn(n-1)/2$ distinct intersection points ζ_{ij} of the form (4.16). In that case each ζ_{ij} is a double root of $D(x)$. We now show that there exist $p(\lambda,x)$ of the form (5.1) for which the polynomial $D(x)$ has $rn(n-1)$ simple (distinct) roots. Let $\delta(p)$ be the discriminant of the polynomial $D(x)$. That is $\delta(p)$ is given by the well known determinental formula

$$(5.9) \quad \delta(p) = \begin{vmatrix} d_0(p) & d_1(p) & \cdots & d_{rn(n-1)}(p) & 0 & \cdots & 0 \\ 0 & d_0(p) & d_1(p) & \cdots & d_{rn(n-1)}(p) & \cdots & 0 \\ \vdots & \ddots & & & & & \\ rn(n-1)d_0(p) & [rn(n-1)-1]d_1(p) & \cdots & 0 & \cdots & & \\ \ddots & & \ddots & & \ddots & & \end{vmatrix}$$

(e.g. Whitney [1972, Appendix IV]).

We claim

<u>Theorem</u> 5.10. Let $\delta(p)$ be a polynomial on $P^{n,r}$ given by (5.9). Then $\delta(p)$ does not vanish identically on $P^{n,r}$. In particular, if $\delta(p) \neq 0$ then $D(x)$ is a polynomial of degree $rn(n-1)$ having $rn(n-1)$ simple roots.

<u>Proof.</u> Let p be a polynomial of the form (4.15) such that $b_i \neq b_j$, for $i \neq j$ and $D(x)$ has exactly $rn(n-1)/2$ double roots of the form (4.16). Let $q \in P^{n,r}$ be in the neighborhood of p. So $D(x,q)$ is a polynomial of degree $rn(n-1)$ with the roots $\xi_1(q),\cdots,\xi_{rn(n-1)}(q)$ continuously depend on q. For $q=p$

$$\xi_{2i-1}(p) = \xi_{2i}(p), \quad i=1,\cdots,rn(n-1)/2, \quad \xi_{2i}(p) \neq \xi_{2j}(p)$$

$$\text{for} \quad 1 \le i < j \le n(n-1)/2.$$

So either we can find q such that $\xi_i(q) \neq \xi_j(q)$ for $1 \leq i < j \leq rn(n-1)$ or we must have $\xi_1(q) \equiv \xi_2(q)$ for all q (renaming the indices if necessary). So suppose that $\xi_1(q) = \xi_2(q)$, and $\xi_1(p) = \xi_2(p) = [(a_1-a_2)/(b_2-b_1)]^{1/r} \neq 0$. Choose

$$q(\lambda,x) = \{[\lambda-(a_1+b_1x^r)][\lambda-(a_2+b_2x^r)] + \varepsilon\xi\} \prod_{3\leq i\leq n} [\lambda-(a_i+b_ix^r)]$$

$$= \{\lambda^2-[a_1+a_2+(b_1+b_2)x^r]\lambda + (a_1+b_1x^r)(a_2+b_2x^r)+\varepsilon\xi \prod_{3\leq i\leq n} [\lambda-(a_i+b_ix^r)]$$

where ε is a small parameter. Then $\xi_1(q)$ and $\xi_2(q)$ must be the roots of the equation

$$[(a_1+a_2) + (b_1+b_2)x^r]^2 = 4[(a_1+b_1x^r)(a_2+b_2x^r) + \varepsilon]$$

which is equivalent to

$$(b_1-b_2)^2x^{2r} + 2(b_1-b_2)(a_1-a_2)x^r + (a_1-a_2)^2 - \varepsilon = 0.$$

But for $\varepsilon \neq 0$ all $2r$ roots of the above equation are distinct. This contradicts our assumption that $\xi_1(q) \equiv \xi_2(q)$. Thus we proved our assertion that there exist $q \in P^{n,r}$ in the neighborhood of a *given* p *such that* $D(x,q)$ has $rn(n-1)$ distinct roots. ∎

We now claim that if $\delta(p) \neq 0$ then $p(\lambda,x)$ is an irreducible polynomial. This follows from the following theorem.

<u>Theorem</u> 5.11. Let $p(\lambda,x)$ be a non-degenerate reducible polynomial of the form (5.1). Suppose that

$$p(\lambda,x) = q_1(\lambda,x)q_2(\lambda,x), \deg q(\lambda,x) = n_1 \geq 1, \quad i=1,2$$

(5.12)

$$q_i = \lambda^{n_i} + \sum_{j=1}^{n_i} q_{ij}(x)\lambda^{n_i-j}, \quad i=1,2.$$

Then $D(x)$ has at most $r[n(n-1)-n_1n_2]$ distinct roots and this number is achievable. In particular any reducible polynomial $p(\lambda,x)$ has at most $r(n-1)^2$ distinct intersection points.

Proof. The equation $p(\lambda,x) = 0$ splits to $q_1(\alpha,x) = 0$, $q_2(\beta,x) = 0$. Thus, if ζ is a finite intersection point for $\lambda(x)$ then one of the following conditions hold

(i) $\alpha_i(\zeta) = \alpha_j(\zeta)$ for $i \neq j$,

(ii) $\beta_i(\zeta) = \beta_j(\zeta)$ for $i \neq j$

(iii) $\alpha_i(\zeta) = \beta_j(\zeta)$.

The possibility (i) means that ζ is an intersection point of $q_1(\lambda,x)$. Therefore we can have at most $rn_1(n_1-1)$ intersection points. In the same way possibility (ii) can happen for at most $rn_2(n_2-1)$ distinct points. To find out how many distinct ζ may satisfy (iii) we look at the resultant of $q_1(\lambda,x)$ and $q_2(\lambda,x)$.

$$R(x) = \prod_{1 \leq i \leq n_1,\, 1 \leq j \leq n_2} (\alpha_i(x) - \beta_j(x)).$$

Since $q_1(\lambda,x)$ and $q_2(\lambda,x)$ are monic it is well known that $R(x)$ is polynomial in the coefficients of λ in $q_1(\lambda,x)$ and $q_2(\lambda,x)$. So $R(x)$ is a polynomial in x. The expansion (5.3) at the infinity yields that $|R(x)| \leq K|x|^{rn_1n_2}$.

So the degree of R is at most rn_1n_2. Therefore there are at most rn_1n_2 distinct ζ satisfying (iii). Altogether we get that $D(x,p)$ has at most the following number of distinct zeros

$$\begin{aligned} r[n_1(n_1-1)+n_2(n_2-1)+n_1n_2] &= r[(n_1+n_2)(n_1+n_2-1)-n_1n_2] \\ &= r[n(n-1)-n_1n_2]. \end{aligned}$$

Let $p(x)$ be of the form (4.15) such that $D(x)$ has exactly $rn(n-1)/2$ double roots. Define

$$p_1(\lambda,x) = \prod_{i=1}^{n_1} [\lambda-(a_i+b_ix^r)], \quad p_2(\lambda,x) = \prod_{i=n_1+1}^{n} [\lambda-(a_i+b_ix^r)] \quad i = n_1+1.$$

Now we can find q_i in the neighborhood of p_i such that $\delta(q_i) \neq 0$, $i=1,2$. Then each q_i will have $rn_i(n_i-1)$ distinct intersection points. As

/

$R(x,p_1,p_2)$ has exactly rn_1n_2 distinct zeros the continuity argument implies that (x,q_1,q_2) will have exactly rn_1n_2 distinct zero. Using the continuity argument again we deduce that all intersection points satisfying (i)-(iii) are pairwise distinct. This shows that $D(x,q_1,q_2)$ has $r[n_1(n_1-1) + n_2(n_2-1)]$ simple roots and rn_1n_2 double roots. Clearly

$$n(n-1) - n_1n_2 \leq (n-1)^2$$

and the equality holds if only either $n_1=1$ or $n_1=n-1$. Hence the maximal number of distinct solutions of $D(x,p) = 0$ is $r(n-1)^2$.

Assume that $p(\lambda,x)$ is degenerate. Then all the intersection points of $p(\lambda,x)$ are the intersection points of

$$q(\lambda,x) = \prod_{i=1}^{k} q_i(\lambda,x) \tag{5.13}$$

where (4.2) is the decomposition of $p(\lambda,x)$ to its irreducible factors. As $\deg q(\lambda,x) < n$ the number of intersection points of q is at most $r(n-2)^2$. Thus any reducible $p(\lambda,x)$ has at most $r(n-1)^2$ intersection points. ■

<u>Corollary</u> 5.14. Let $p(\lambda,x)$ be of the form (5.1). If $\delta(p) \neq 0$ then $p(\lambda,x)$ is irreducible.

We now study the set of reducible and degenerate polynomials $p(\lambda,x)$ of the form (5.1) in the coefficient space $P^{n,r}$. Denote by M_{rd} and M_{dg} the subsets of $P^{n,r}$ corresponding to reducible and degenerated polynomials respectively. Clearly $M_{dg} \subseteq M_{rd}$.

<u>Definition</u> 5.15. Let $M(n_1,n_2)$ be a subset of $P^{n,r}$ which corresponds to reducible polynomial $p(\lambda,x)$ of the form (5.12). Let $N(n_1,n_2)$ be a subset of $P^{n,r}$ which corresponds to degenerate polynomials of the form

$$q(\lambda,x) = q_1(\lambda,x)^2 q_3(\lambda,x), \ \deg q_1 \geq 1 \tag{5.16}$$

where $q_1(\lambda,x)$ is monic in λ (so $q_3 \equiv 1$ if $\deg q_1(\lambda,1) = n/2$).

Clearly

$$(5.17)\qquad M_{rd} = \bigcup_{1\le n_1\le[\frac{n}{2}]} M(n_1,n-n_1),\quad M_{dg} = \bigcup_{1\le n_1\le[\frac{n}{2}]} N(n_1,n-2n_1)$$

$$(5.18)\qquad N(n_1,n-2n_1) \subseteq M(n_1,n-n_1).$$

Theorem 5.19. The sets $M(n_1,n-n_1)$ and $N(n_1,n-2n_1)$ are irreducible algebraic varieties in $P^{n,r}$ of the dimensions

$$\dim M(n_1,n-n_1) = \{[(n+1)r+2]n - 2n_1(n-n_1)r\}/2$$

(5.20)

$$\dim N(n_1,n-2n_1) = \{[(n+1)r+2]n-n_1[r(4n-5n_1+1) + 2]\}/2.$$

Proof. The equality (5.12) can be repesented by a map

$$(5.21)\qquad \theta : P^{n_1,r} \times P^{n_2,r} \to P^{n,r},\quad n_1+n_2 = n$$

where q_i is a coefficient vector in $P^{n_i,r}$. Clearly θ is a polynomial map. The image of θ is exactly $M(n_1,n_2)$. According to Section 3 the closure of $M(n_1,n_2)$ is an algebraic variety in $P^{n,r}$. We now show that $M(n_1,n_2)$ is a closed manifold. Assume that we have a sequence of polynomials $p_i(\lambda,x)$ of the form (5.1) such that $p_i(\lambda,x) = q_{1i}(\lambda,x)q_{2i}(\lambda,x)$, $\deg q_{ji}(\lambda,1) = n_j$, $j = 1,2$ and each $q_{ji}(\lambda,x)$ is monic in λ. Let $\alpha_1^{(i)}(x),\cdots,\alpha_{n_1}^{(i)}(x)$ and $\beta_1^{(i)}(x),\cdots,\beta_{n_2}^{(i)}(x)$ be the roots of $q_{1i}(\lambda,x) = 0$ and $q_{2i}(\lambda,x) = 0$ respectively. Assume that

$$\lim_{i\to\infty} p_i(\lambda,x) = q(\lambda,x).$$

Then it is possible to find a subsequence $\{i_k\}$ such that

$$\lim_{k\to\infty} \alpha_x^{(i_k)}(x) = \alpha_s(x),\quad \lim_{k\to\infty} \beta_t^{(i_k)}(x) = \beta_t(x),\quad 1\le s\le n_1,\ 1\le t\le n_2.$$

Hence (5.12) holds where

$$q_1(\lambda,x) = \prod_{s=1}^{n_1} (\lambda-\alpha_s(x)),\quad q_2(\lambda,x) = \prod_{t=1}^{n_2} (\lambda-\beta_t(x)).$$

So $M(n_1,n_2)$ is closed. The decomposition (5.17) yields that M_{rd} is

closed. So if $p \in P^{n,r}$ corresponds to an irreducible polynomial there exists a neighborhood of p corresponding entirely to irreducible polynomials. We now prove the first equality in (5.20). Pick up irreducible polynomials $q_j(\lambda,x)$ of degree n_j, $j=1,2$, such that $q_1 \neq q_2$. We claim that θ is a local homomorphism in the neighborhood of (q_1,q_2). Indeed suppose that

$$p(\lambda,x) = u_1(\lambda,x)u_2(\lambda,x) = v_1(\lambda,x)v_2(\lambda,x)$$

where (u_1,u_2) and (v_1,v_2) are in the neighborhood of (q_1,q_2). So u_i and v_i are irreducible. Since $C[\lambda,x]$ is a unique factorization domain and u_i and v_i are monic we get either $u_i = v_i$ or $u_1 = v_2$. The assumption that $q_1 \neq q_2$ implies that $u_1 \neq v_2$ so $u_1 = v_1$ and $u_2 = v_2$. Whence θ is a local homomorphism in the neighborhood of (q_1,q_2). Thus

$$\dim M(n_1,n_2) = \{[(n_1+1)r+2]n_1 + [(n_2+1)r+2]n_2\}/2$$

which establishes the first inequality in (5.20). It is left to show that $M(n_1,n_2)$ is an irreducible variety. This follows easily from the fact that θ is a local homeomorphism in the neighborhood of (q_1,q_2) for $q_1 \neq q$.

The assertions about $M(n_1,2n-2n_1)$ can be proven in the analogous way. ∎

As

$$\{[(n+1)r+2]n-2n_1(n-n_1)r\}/2 < \{[(n+1)r+2]n-2(n-1)r\}/2$$

$$\{[(n+1)r+2]n-n_1[r(4n-5n_1+1)+2]\}/2 < \{[(n+1)r+2]n-[4r(n-1)+2]\}/2$$

from the identities (5.17) we get

<u>Corollary</u> 5.22. The sets M_{rd} and M_{dg} are algebraic varieties in $P^{n,r}$ having the following codimensions

(5.23) $\qquad \operatorname{codim} M_{rd} = (n-1)r, \quad \operatorname{codim} M_{dg} = 2r(n-1)+1.$

Moreover M_{dg} is an algebraic subvariety of M_{rd}.

6. Irreducible pencils.

Let $A,B \in M_n$. With the pencil $A+xB$ we associate its characteristic polynomial

$$p(\lambda,x,A,B) = |\lambda I-(A+xB)| = \lambda^n + \sum_{i=1}^{n} p_i(x)\lambda^{n-i},$$

(6.1)

$$p_i(x) = \sum_{j=0}^{i} p_{ij}(A,B)x^j, \qquad i = 1,\cdots,n.$$

Clearly $p_{ij}(A,B)$ is a polynomial of degree $(i-j)$ in the entries of A and of degree j in the entries of B. So $p_{ij}(A,B)$ is a polynomial of (total) degree i. Let (D,E) be a similar pair in (A,B). Obviously, the pencil $D+xE$ has the same characteristic polynomial. So each $p_{ij}(A,B)$ are invariant polynomials under the action (TAT^{-1},TBT^{-1}). In fact, for $n=2$, it is easy to show that the ring generated by $\phi_i(A,B)$, $j=0,\cdots,i$, $i=1,2$ is equal to the ring generated by $\phi_i(A,B)$, $i=1,\cdots,5$, given by (2.1). We now give the explicit expression for $p_{ij}(A,B)$. We use the notation of Section 1. Then the coefficients of the characteristic polynomial of $A+xB$ are given by

$$p_1(x) = \sum_{\alpha\in Q_{i,n}} (A+xB)[\alpha|\alpha],$$

$$p_{ij}(A,B) = \sum (-1)^{\varepsilon(\beta_1,\beta_2,\gamma_1,\gamma_2)} |A[\beta_1|\gamma_1]|\,|B[\beta_2|\gamma_2]|, \quad j=1,\cdots,i-1$$

(6.2) $$\beta_1,\gamma_1 \in Q_{(i-j),n}, \beta_2,\gamma_2 \in Q_{j,n}, \ \beta_1 \quad \beta_2 = \gamma_1 \quad \gamma_2 = \phi,$$

$$\beta_1 \quad \beta_2 = \gamma_1 \quad \gamma_2 = \alpha$$

$$p_{i0}(A,B) = \sum_{\alpha\in Q_{i,n}} |A[\alpha|\alpha]|, \quad p_{ii}(A,B) = \sum_{\alpha\in Q_{i,n}} |B[\alpha|\alpha]|, \quad i=1,\cdots,n.$$

Here $\varepsilon(\beta_1,\beta_2,\gamma_1,\gamma_2) = \pm 1$ and this function is completely determined by β_1,β_2,γ_1 and γ_2 which satisfy the requirements (6.2). Also for $\alpha \in Q_{k,n}$, $\beta \in Q_{\ell,n}$ we denote by $\alpha \cap \beta$ the common subsequence of α and β. If $\alpha \cap \beta = \phi$ then $\alpha \cup \beta$ denotes the strictly increasing sequence generated by the elements of α and β.

The fact that with each pair (A,B) we associate a polynomial of the form (6.1) can be put formally in terms of a polynomial map

(6.3)
$$p : M_n \times M_n \to \mathbb{C}^{(n+3)n/2} = P^{n,1}$$
$$p(A,B) = (p_{10}(A,B),\cdots,p_{nn}(A,B)).$$

We shall show in the sequel that p is a regular map (note that $M_n \equiv \mathbb{C}^{n^2}$).

First we need

Definition 6.4. The pencil $A+xB$ (the pair (A,B)) is called reducible (degenerated) if the characteristic polynomial (6.1) is reducible (degenerated). Otherwise the pencil (the pair) is called irreducible (non-degenerate). The pencil $A+xB$ (the pair (A,B)) is called symmetric (real) if A and B are symmetric (real).

We now show that there exist irreducible pencils. More precisely we have

Theorem 6.5. Let $\delta(p)$ be a polynomial defined on the coefficient space $P^{n,1}$ by (5.9) $(r=1)$. Then $\delta(p(A,B))$ is a non-trivial polynomial on $M_n \times M_n$. More precisely, there exist a real symmetric pair (A,B) such that $\delta(p(A,B)) \neq 0$.

Proof. Our proof is very close to the proof of Theorem 5.10. So we point out only the additional arguments we have to use. Choose A and B to be real diagonal $A = \text{diag}\{a_1,\cdots,a_n\}$, $B = \text{diag}\{b_1,\cdots,b_n\}$. Then the characteristic polynomial of $A+xB$ is of the form (4.15) with $r=1$. We choose a_i and b_j such that $p(\lambda,x)$ will have exactly $n(n-1)/2$ distinct intersection points. That is $D(x)$ given by (4.17) $(r=1)$ has exactly $n(n-1)/2$ double roots. Suppose that $\delta(p(D,E)) \equiv 0$. Then $D(x,D,E)$ has to have a fixed double root let us say $\xi_1(D,E) = \xi_2(D,E)$ such that

$$\xi_1(A,B) = \xi_2(A,B) = (a_1-a_2)/(b_2-b_1).$$

Choose $E = B$ and let D be a block diagonal matrix

$$D = \operatorname{diag}\{D_1, \cdots, D_{n-1}\}$$

where D_{i+1} is 1×1 matrix $\{a_{i+2}\}$ $i=1,\cdots,n-2$ and

$$D_1 = \begin{pmatrix} a_1 & \varepsilon \\ \varepsilon & a_2 \end{pmatrix}.$$

As in the proof of Theorem 5.10 we deduce that $\xi_1(D,E) \neq \xi_2(D,E)$ for $\varepsilon \neq 0$. This shows that $\delta(p(D,E))$ cannot vanish identically on all real symmetric pairs. ■

According to Corollary 5.14 if $\delta(p(A,B)) \neq 0$ then the characteristic polynomial of $A+xB$ is irreducible. So "most" of the pencils $A+xB$ are irreducible. Consider a pair (A,B). Assume that A has n distinct eigenvalues. Then A is similar to a diagonal matrix

(6.6) $$D = \operatorname{diag}\{d_1, \cdots, d_n\}.$$

So the orbit of (A,B) *contains a pair of the form* (D,E). Clearly E is unique up to a diagonal similarity XEX^{-1} (X is diagonal). Let $E = (e_{ij})_1^n$. Denote

(6.7) $$E_+ = (|e_{ij}|)_1^n.$$

<u>Definition</u> 6.8. The matrix E is called irreducible if the nonnegative matrix E_+ is irreducible. That is, all the entries of $(I+E_+)^{n-1}$ are strictly positive. (In other terms the graph defined by E_+ is connected).

Clearly the notion of irreducibility remains invariant under the diagonal similarity. In what follows we need a criterion for the diagonal similarity (see Engel-Schneider [1973]).

<u>Theorem</u> 6.9. Let $E,F \in M_n$. Assume E is irreducible. Then E and F are diagonally similar if and only if

$$e_{ii} = f_{ii},\ i=1,\cdots,n,\quad e_{i_1 i_2} e_{i_2 i_3} \cdots e_{i_{k-1} i_k} e_{i_k i_1} =$$

(6.10)

$$= f_{i_1 i_2} f_{i_2 i_3} \cdots f_{i_{k-1} i_k} f_{i_k i_1},\quad 1 \le i_j \le n,\ j=1,\cdots,k,\ 2 \le k \le n.$$

On the other hand if E_+ is reducible the equalities (6.10) do not imply the diagonal similarity of E and F. We now give a set of invariant polynomials in $[M_n \times M_n]$ whose values determine uniquely the orbit of (A,B) in case that (A,B) is an irreducible pair.

Theorem 6.11. Let $A,B \in M_n$. The polynomials

(6.12) $$\mathrm{tr}(A^{i_1}B^{j_1},\cdots,A^{i_m}B^{j_m}),\quad m = n(n-1),\quad 0 \le i_k,j_k \le 1,\ k=1,\cdots,m$$

are invariant polynomials under the simultaneous similarity. Moreover two irreducible pairs (A,B) and (D,E) are simultaneously similar if and only if the above polynomials have the same values on these pairs.

Proof The fact that the polynomials (6.12) are invariant with respect to the action of GL_n is obvious. Suppose next that A has n distinct eigenvalues. Then the orbit of (A,B) contains a matrix (D,E) where D is a diagonal matrix with pairwise distinct diagonal entries. The entries of D are known since we are given $\mathrm{tr}(A^i)$, $i=1,\cdots,n$. We claim that E is irreducible. Otherwise (e.g. Gantmacher [1959]) there exists a permutation matrix P such that

(6.13) $$PEP^t = \begin{pmatrix} F_{11} & F_{12} \\ & \\ 0 & F_{22} \end{pmatrix}$$

Also PDP^t is a diagonal matrix of the form $\mathrm{diag}\{D_1,D_2\}$. But then

$$|\lambda I-(A+xB) = |\lambda I_1-(D_1+xF_{11})|\,|\lambda I_2-(D_2+xF_{22})|$$

which contradicts the assumption that $A+xB$ is an irreducible pencil.

According to Theorem 6.9 E is determined up to a diagonal similarity provided we can compute the left-hand side of the equality (6.10).

Let

$$D_i = \operatorname{diag}\{\delta_{i1},\cdots,\delta_{in}\}.$$

Then a straightforward calculation shows that

$$(6.14)\qquad \begin{aligned} &e_{ii} = \operatorname{tr}(D_i E) \\ &e_{i_1 i_2} e_{i_2 i_3} \cdots e_{i_{k-1} i_k} e_{i_k i_1} = \operatorname{tr}(D_{i_1} E D_{i_2}, \cdots, D_{i_k} E). \end{aligned}$$

Since all the eigenvalues of D are pairwise distinct we have

$$D_i = p_i(D), \quad p_i(\lambda) = \prod_{j \neq i} \frac{(\lambda-\lambda_j)}{(\lambda_i-\lambda_j)}, \qquad i=1,\cdots,n.$$

So the right hand side of (6.14) contains expressions of the form

$$(6.15)\qquad \operatorname{tr}(D^{\ell_1} E D^{\ell_2} E, \cdots, D^{\ell_k} E), \quad 0 \le \ell_i \le n-1, \quad i=1,\cdots,k, \quad 1 \le k \le n-1.$$

Substituting A for D and B for E we realize that the above expression are included in the expressions appearing in (6.12). This proves the theorem in case that A has a distinct eigenvlaue. In the general case let $A_0 = A+x_0B$.

Clearly we can choose x_0 such that A_0 has a distinct eigenvalue. Then in the expression (6.15) we have to substitute $A+x_0B$ for D and B for E. A straightforward calculation also shows that all possible expressions in terms of A and B which may appear in (6.15) are listed in (6.12). The proof of the theorem is completed. ∎

In fact we proved a more precise statement.

<u>Theorem</u> 6.16. Let A be a diagonal matrix with pairwise distinct elements and B is irreducible. Then the orbit of (A,B) is determined uniquely by the values (6.12).

The distinction between Theorems 6.11 and 6.16 becomes apparent for $n \geq 3$. Indeed, if A is a diagonal matrix with pairwise distinct elements and $A+xB$ is irreducible then B is an irreducible matrix. However, for $n \geq 3$ there exists A and B satisfying the assumptions of Theorem 6.16

such that $A+xB$ is a reducible pencil (e.g. Friedland-Simon [1981]). In fact, the assumption that any pencil satisfying the assumption of Theorem 6.16 is irreducible, implies a false conjecture in (Avron-Simon [1974], Friedland-Simon [1981]).

Thus we proved

<u>Theorem</u> 6.17. The values of the invariant functions (6.12) separates all orbits (A,B) if A and B does not have a common subspace and $A+xB$ is not a degenerated pencil.

Let D be a diagonal matrix with pairwise distinct diagonal entries. Consider a matrix $C = (c_{ij})^n$, such that $c_{ij} \neq 0$, $i=2,\cdots,n$. Then we can find a diagonal matrix X such that $E = XCX^{-1} = (e_{ij})_1^n$ and $e_{1i} = 1$, $i = 2,\cdots,n$.

In fact X is unique up to a multiplication by $\lambda \neq 0$. That is in "general" we can find a unique pair

$$(D,E),\ D = \mathrm{diag}\{d_1,\cdots,d_n\},\ d_i \neq d_j \ \text{ for } i \neq j, \tag{6.18}$$
$$E = (e_{ij})_1^n,\ e_{1i} = 1,\ i=2.\cdots,n,$$

in the orbit of (A,B). That is if $\mathrm{orb}(A_i,B_i)$ contains pairs of (D_i,E_i) of the form (6.18) then $\mathrm{orb}(A_1,B_1) = \mathrm{orb}(A_2,B_2)$ if and only if $(D_1,E_1) = (D_2,E_2)$. That shows that the "generic" orbits are parametrized by n^2+1 parameters. If the entries of the matrix E given in (6.18) are all distinct from zero then $\mathrm{orb}(D,E)$ is fixed by the values of polynomials (6.12) in view of Theorem 6.11. So the ring of invariant polynomials $[M_n \times M_n]$ separates between the "generic" orbits. That is

<u>Theorem</u> 6.19. The transendence dimension of the ring of invariant polynomial $[M_n \times M_n]^G$ is n^2+1.

We were not able to find a simple transendence basis in $[M_n \times M_n]^G$ for a general n.

We now show that the functions (6.12) generate the field of invariants $(M_n \times M_n)^G$.

<u>Theorem</u> 6.20. The invariant polynomials given in (6.12) generate the field of rational invariant functions $(M_n \times M_n)^G$.

<u>Proof.</u> Let $\varphi_1,\cdots,\varphi_k$ $(k=2^{n(n-1)})$ be the invariant polynomials given in (6.12). In Theorem 6.11 we showed that the values of these polynomials determine the orbit of an irreducible pair (A,B). Since the set of irreducible pairs (A,B) form an open (algebraic) set in $M_n \times M_n$ it follows that $\varphi_1,\cdots,\varphi_k$ form a transcendal basis in $(M_n \times M_n)^G$. That is any $\theta \in (M_n \times M_n)^G$ satisfied the equation

$$\theta^m + \sum_{i=0}^{m} \rho_i(\varphi_1,\cdots,\varphi_k)\theta^{m-i} = 0. \tag{6.21}$$

Let $\varphi : M_n \times M_n \to \mathbb{C}^k$ be the polynomial map given by $\varphi = (\varphi_1,\cdots,\varphi_k)$. Denote by V the closure of $\varphi(M_n \times M_n)$. So θ is an algebraic function on V. On the other hand let V^0 be the set of all $\varphi(A,B)$ such that (A,B) is an irreducible pair. So V^0 is an open set in V. Since for any $\varphi(A,B) \in V^0$, $\varphi(A,B)$ determines the $\mathrm{orb}(A,B)$ we see that $\theta(\varphi(A,B))$ is defined uniquely. So the minimal m in (6.22) can be chosen to be 1. That is θ is rational in $\varphi_1,\cdots,\varphi_k$. Hence $\varphi_1,\cdots,\varphi_k$ generates the field $(M_n \times M_n)^G$. ∎

In conclusion we note that according to Processi [1976] the invariant polynomials (6.12) generate the ring of invariant polynomials $[M_n \times M_n]^G$.

7. Symmetric pairs.

In this section we are going to consider (complex) symmetric pairs of matrices (A,B). Two pairs of symmetric matrices (A,B) and (D,E) are called orthogonally similar if there exists (complex) orthogonal matrix O such that

$$D = O^t AO, \quad E = O^t BO. \tag{7.1}$$

Denote by $S_n(S_n(\mathbf{R}))$ the set of $n \times n$ complex (real) valued matrices. So the complex (real) orthogonal group $O_n(O_n(\mathbf{R}))$ acts in the above form on $S_n \times S_n(S_n(R) \times S_n(R))$. Suppose that A is symmetric and has n distinct eigenvalues. Then there exists an orthogonal matrix O such that $D = O^t AO$ is a diagonal matrix. Thus the orbit of (A,B) contains a symmetric pair (D,E). We claim that E is fixed up to an action of the following finite orthogonal group,

$$DO_n = \{D | D = \operatorname{diag}\{1,d_2,\cdots,d_n\}, \quad d_i \leq 1, \quad i = 2,\cdots,n\}. \tag{7.2}$$

Lemma 7.3. Let $E,F \in S_n$. Assume that E and F are diagonally similar. Then there exists $D \in DO_n$ such that $F = DED$.

Proof. Suppose that $F = DED^{-1}$, $D = \operatorname{diag}\{d_1,\cdots,d_n\}$. Assume that $e_{ij} \neq 0$, then

$$f_{ij} = d_i e_{ij} d_j^{-1} = f_{ji} = d_j e_{ji} d_i^{-1}.$$

That is $d_1^2 = d_j^2$. Suppose that E is irreducible. We then deduce $d_1^2 = \cdots = d_n^2$. Clearly we can choose $d_1 = 1$. This shows that $D \in OD_n$. Suppose that E is reducible. Since E is symmetric there exists a permutation matrix P such that

$$P^t EP = \operatorname{diag}\{E_1,\cdots,E_k\}$$

where each E_j is irreducible. But then

$$P^t FP = \operatorname{diag}\{F_1,\cdots,F_k\}, \quad P^t FP = P^t DP(P^t EP)P^t D^{-1}P$$

So each F_i is a diagonally similar to an E_i. Now use the above arqument to show that D can be chosen to be in DO_n. ∎

Thus if $A \in S_n$ has n distinct eigenvalues a pair (D,E) essentially parametrize (up to the action of DO_n) to orbit of (A,B). Hence "most of the orbits" are parametrized by $n + n(n+1)/2$ parameters. That is the transcendence degree of $[S_n \times S_n]^0$ is $(n+3)n/2$.

It is easy to find a transcendence basis in $[S_n \times S_n]^0$.

Theorem 7.4. The polynomials

$$\mathrm{tr}(A^i),\ i = 1,\cdots,n,\quad \mathrm{tr}(A^iBA^jB),\quad 0 \le i < j \le n-1 \tag{7.5}$$

from a transcendence basis in $[S_n \times S_n]^0$. Moreover, if A has n distinct eigenvalues then the value of these polynomials determine at most $2^{(n^2-n+2)/2}$ distinct orbits.

Proof. First note that $\mathrm{tr}(A^iBA^jB] = \mathrm{tr}(BA^jBA^i)$, for any pair (i,j). As $\mathrm{tr}(A^i)$, $i = 1,\cdots,n$ are given we know the characteristic polynomial of A. Using the values of the polynomials in (7.5) and the remark above we can compute any polynomial of the form $\mathrm{tr}(A^iBA^jB)$. Assume next that A has n distinct eigenvalues. So the orbit of (A,B) contains a symmetric pair (D,E) where D is a diagonal matrix with pairwise distinct diagonal entries. The entries of D are known since we are given $\mathrm{tr}(A^i)$, $i = 1,\cdots,n$. Using the arguments of the proof of Theorem (6.11) we deduce that we can compute the products e_{ij}, e_{ji} for all $1 \le i, j \le n$ in terms of the given polynomials (7.5). As E is symmetric it follows that we know the values of e_{ij}^2. So each e_{ij} is fixed up to ± 1. There are at most $2^{n(n+1)/2}$ different matrices E. However (D,E) and (D,XEX) are in the same orbit for any $X \in DO_n$. Thus we have at most $2^{(n+1)n/2-(n-1)}$ distinct orbits corresponding to the values of the polynomials in (7.5). In fact, if all e_{ij} are different from zero then the knowledge of D and all e_{ij}^2 qives

rise to exactly $2^{(n^2-n+2)/2}$ distinct orbits, provided that D has pairwise distinct entries. Since the orbit space of symmetric pairs (A,B) is parametrized by $(n+3)n/2$ parameters we see that the polynomials (7.5) form a transcendence basis in $[S_n \times S_n]^0$. ■

Combine the arguments of the proofs of Theorems 7.4 and 6.17 to get

Theorem 7.6. Let (A,B) and (D,E) be non-degenerate symmetric paris. Assume that neither A and B nor D and E have a common subspace. Then these pairs are orthogonally similar if and only if the polynomials given by (6.12) have the same values on these pairs. ■

We now classify the orbits of 2×2 symmetric pairs under the orthogonal similarity. Let U and V be subvarieties of $S_2 \times S_2$ given by (2.3) and (2.8) respectively. Suppose that (A,B) lies in U but $(A,B) \notin V$. Then the arguments of Theorem 2.9 show that (A,B) are orthogonally similar to a pair of diagonal matrices (D,F). Whence it follows

Theorem 7.7. Let $(A,B) \in S_2 \times S_2$. Assume that $(A,B) \notin V$ (i.e. (A,B) is non-degenerate) then the values of the polynomials $\varphi_i(A,B)$, $1 \leq i \leq 5$ determines a unique orbit under the action of O_2.

Suppose that $C \in S_2$ and C has one multiple eigenvalue. It is easy to show that either C is of the form

$$(7.8) \qquad C(\lambda,a) = \begin{pmatrix} \lambda+a & ai \\ ai & \lambda-a \end{pmatrix}$$

or $C = C(\lambda,-a)$. Moreover if $a \neq 0$ then $C(\lambda,a)$ is orthogonally similar to any $C(\lambda,b)$ or $C(\lambda,-b)$, $b \neq 0$. Thus if $(A,B) \in V$ we have that

$$A = C(\lambda,\pm a), \quad B = C(\mu,\pm b).$$

Therefore the function $\gamma(A,B)$ given by (2.24) is equal to $\pm a/b$ and is a nontrivial function in $(V)^{O_2}$. So Theorem 2.23 applies to $S_2 \times S_2$.

Theorem 7.9. Let $(A,B) \in V \quad S_2 \times S_2$. Then $\gamma(A,B) = a_{12}/b_{12}$ belongs to $(V)^{O_2}$. If either $\gamma(A,B)$ or $1/\gamma(A,B)$ is defined on the orbit of (A,B) then the values of $\mathrm{tr}(A)$, $\mathrm{tr}(B)$ and $\gamma(A,B)(1/\gamma(A,B))$ determine a unique orbit in V under the action of O_2. Otherwise the orbit of (A,B) consists of one point $\left(\frac{\mathrm{tr}(A)}{2} I, \frac{\mathrm{tr}(B)}{2} I\right)$.

The disadvantage of the transcendence basis (7.5) of $[S_n \times S_n]^0$ is that these polynomials are not symmetric with respect to A and B. The natural candidate for symmetric basis in $[S_n \times S_n]^0$ are the coefficients of the characteristic polynomial of the pencil $A+xB$. Indeed, the map (6.3) restricted to $S_n \times S_n$ yields

$$(7.10) \qquad p : S_n \times S_n \to C^{(n+3)n/2} = P^{n,1}.$$

Thus, if $p_{10}(A,B), \cdots, p_{nn}(A,B)$ are algebraically independent it follows that these polynomials form a transcendenal basis in $[S_n \times S_n]^0$ since the transcendence degree of $[S_n \times S_n]^0$ is $(n+3)n/2$. Also the polynomials $p_{ij}(A,B)$ do exhibit symmetricity in A and B since

$$(7.11) \quad p_{ij}(A,B) = p_{i(j-i)}(B,A), \quad 0 \leq j \leq i, \quad 0 \leq i \leq n.$$

Clearly, $p_{ij}(A,B)$, $0 \leq j \leq i$, $0 \leq i \leq n$ form a transcendence basis in $[S_n \times S_n]$ if and only if the map (7.10) is regular. If we can show that p is onto map then of course p is regular. Suppose that we showed that p is an onto map for $k \leq n-1$. Thus if $p(\lambda,x)$ is a reducible polynomial of the form (5.1) then there exists a pair of symmetric block diagonal matrices (A,B),

$$A = \mathrm{diag}\{A_1, A_2\}, \quad B = \mathrm{diag}\{B_1, B_2\}$$

such that

$$|\lambda I - (A+xB)| = p(\lambda,x).$$

So, in order to prove that p is an onto map, it is enough to show that any irreducible polynomial $p(\lambda,x)$ of the form (5.1) is a characteristic of some

symmetric pencil $A+xB$. By making a transformation $x = y + x_0$ it is enough to consider irreducible polynomials with the property

$$(7.12)\qquad p(\lambda,0) = \prod_{i=1}^{n} (\lambda-d_i), \quad d_i \neq d_j \quad \text{for} \quad i \neq j.$$

If $|\lambda I - (A+xB)| = p(\lambda,x)$ then we find a symmetric pair (D,E) $D = \mathrm{diag}\{d_1,\cdots,d_n\}$ such that $|\lambda I - (D+xE)| = p(\lambda,x)$.

Thus, if p is onto map then the map

$$(7.13)\qquad p(D,\cdot) : S^n \to P^{n,1}$$

$$P(D,E) = (p_{11}(D,E),\cdots,p_{n-1}(D,E),\cdots,p_{nn}(D,E))$$

is an onto map. Clearly, $p_{ji}(D,E)$ is a homogeneous polynomial of degree i in the entries of E. According to Theorem 3.20 if the system

$$(7.14)\qquad p_{ij}(D,E) = 0, \quad j = 1,\cdots,i, \quad i = 1,\cdots,n$$

has a unique solution $E = 0$ then the map (7.13) is an onto map. The system (7.14) is equivalent to the assertion that the pencil $D+xE$ has constant eigenvalues (spectrum) $d_1,\cdots,d_n$. The main result of the next section is

Theorem 7.15. Let D be a diagonal marix with pairwise distinct diagonal entries. Assume that the symmetric pencil $D+xE$ has a constant spectrum. Then $E = 0$ for $1 \leq n \leq 4$. For $n \geq 5$ and a given D there exist nontrivial E satisfying (7.14).

8. Polynomial matrices with a constant spectrum.

In what follows we adopt the following notation. By $M_n(\mathbf{C}[x])$ we denote $n\times n$ matrices with polynomial entries. That is if $A(x) \in M_n(\mathbf{C}[x])$ then

$$(8.1) \qquad A(x) = (a_{ij}(x))_1^n = \sum_{j=0}^{m} A_j x^j, \quad A_j \in M_n.$$

Let $GL_n(\mathbf{C}[x])$ be the general linear group over the ring $\mathbf{C}[x]$. That is $U(x) \in M_n(\mathbf{C}[x])$ is in $GL_n(\mathbf{C}[x])$ if and only if there exists $V(x) \in M_n(\mathbf{C}[x])$ such that $U(x)V(x) = I$. Denote by $O_n(\mathbf{C}[x])$ the orthogonal subgroup of $M_n(\mathbf{C}[x])$. That is $O_n(\mathbf{C}[x])$ consists of all $U(x)$ such that $U(x)U^t(x) = I$. Denote by $S_n(\mathbf{C}[x])$ and $A_n(\mathbf{C}[x])$ the subsets of symmetric and skew symmetric matrices in $M_n(\mathbf{C}[x]$ respectively. Here A_n denotes the set of skew symmetric matrices $(A^t = -A)$ in M_n. Let $F(x) \in M_n(\mathbf{C}[x])$ and consider the differential equation

$$(8.2) \qquad \frac{dU}{dx} = UF(x)$$

with the initial condition $U(0) = I$. Then $U(x)$ is invertible for each x. In fact it is easy to see that $V = U^{-1}(x)$ satisfies the equation

$$(8.3) \qquad \frac{dV}{dx} = -F(x)V$$

with $V(0) = I$. Clearly, $U(x)$ is orthogonal for each x if and only if $F(x)$ is a skew symmetric matrix. However, usually $U(x)$ will not belong to $M_n(\mathbf{C}[x])$.

Theorem 8.4. Let $U(x) \in GL_n(\mathbf{C}[x])$. Then $U(x)$ satisfies the differential equation (8.2) with $F(x) \in M_n(\mathbf{C}[x])$

$$(8.5) \qquad F(x) = \sum_{j=0}^{\ell} F_j x^j.$$

Moreover F_ℓ must be nilpotent.

Proof. Define

$$F(x) = U^{-1}\frac{dU}{dx}.$$

So $F(x) \in M_n(\mathbf{C}[x])$ and U satisfies (8.2).

Suppose that F is of the form (8.5). Then around $x = \infty$ the leading part of (8.2) reduces to

$$\frac{dW}{dx} = x^{\ell} W F_{\ell} .$$

Thus around $x = \infty$ the solution $U(x)$ behaves as $U(\xi)\exp(F_{\ell} x^{\ell+1}/\ell+1)$ $(|\xi| >> 1)$.

If F_{ℓ} is not nilpotent some entries of $U(x)$ behave at infinity as $e^{\rho x^{\ell+1}}$ $(\rho \neq 0)$. Since $U(x)$ is a polynomial in x we deduce that F_{ℓ} is nilpotent. (See for example Lutz [1967] for the precise version of this result). ■

Let E be a nilpotent matrix then

(8.6) $\qquad U(x) = e^{p(x)E}, \quad p(x) \in C[x].$

belongs to $GL_n(\mathbf{C}[x])$. In fact we have

<u>Theorem</u> 8.7 The group $GL_n(\mathbf{C}[x])$ is generated by GL_n and the matrices of the form (8.6).

<u>Proof.</u> Let $U(x) \in GL_n(\mathbf{C}[x])$. It is well known (e.g. Gantmacher [1959, ch. 6]) that $U(x)$ can be brought by the elementary operation to its Smith normal form which is the identity matrix I. Each such elementary operation is carried out by multiplications from left or right by the following types of martices $A(x)$

(i) $\qquad A(x) \equiv A \in GL_n,$

(ii) $\qquad A(x) = I + p(x)E_{ij} = \exp(p(x)E_{ij}), \quad i \neq j,$

where E_{ij} is the matrix whose (i,j) entry is 1 and all other entries vanish. So $E_{ij}^2 = 0$. This proves the theorem. ■

Suppose that $E \in A_n$ (the set of skew symmetric matrices) and is nilpotent. Then the matrix $U(x)$ given by (8.6) is orthogonal. Clearly for $n=2$ the only skew symmetric nilpotent matrix is the zero matrix. Therefore, if $U(x) \in O_2(\mathbf{C}[x])$ the corresponding $F(x)$ in (8.2) must be zero

matrix in the virtue of Theorem 8.4. That is

$$(8.8) \qquad O_2(\mathbf{C}[x]) = O_2.$$

This also follows from the simple fact that the squares of the elements in each row of an orthogonal matrix $U(x)$ sum to 1. Indeed the equality

$$1 = u^2(x) + v^2(x) = (u(x) + iv(x))(u(x) - iv(x))$$

imply that $u(x)$ and $v(x)$ are constants if $u(x)$ and $v(x)$ are polynomials. For $n \geq 3$ there are non-zero skew symmetric matrices. For example

$$(8.9) \qquad E = \mathrm{diag}(H,0), \quad H = \begin{pmatrix} 0 & 1+i & 0 \\ -1-i & 0 & -1+i \\ 0 & 1-i & 0 \end{pmatrix}.$$

So we pose an obvious problem

Problem 8.10. Is $O_n(\mathbf{C}[x])$ generated by O_n and the matrices of the form (8.6) $(E^t = -E)$ for $n \geq 3$?

Let $A(x) \in M_n(\mathbf{C}[x])$. $A(x)$ is called rank 1 matrix if $A(x) \not\equiv 0$ and all 2×2 minors of A vanish identically. Suppose that $A \in M_n$ is rank one matrix. Then $A = (u_i v_j)_1^n$, where $u = (u_1,\cdots,u_n)^t$ and $v = (v_1,\cdots,v_n)^t$ span the ranges of A and A^t respectively. For a rank 1 matrix $A(x)$ we have a similar result.

Theorem 8.11. Let $A(x) \in M_n(\mathbf{C}[x])$ be rank 1 matrix. Then there exists polynomials $a(x)$, $u_1(x),\cdots,u_n(x)$, $v_1(x),\cdots,v_n(x)$ such that

$$(8.12) \qquad A(x) = (a_{ij}(x))_1^n, \; a_{ij}(x) = a(x)u_i(x)v_j(x), \; i,j = 1,\cdots,n$$

and u_i, $i = 1,\cdots,n$ $(v_i, i = 1,\cdots,n)$ do not have common zeros. Moreover if $A(x)$ is symmetric it is possible to choose $u_i(x) = v_i(x)$, $i = 1,\cdots,n$.

Proof. Choose ξ and constant vectors α and β such that $A(\xi)\alpha \neq 0$, $A^t(\xi)\beta \neq 0$. So $A(x)\alpha = (a_1(x)(u_1(x),\cdots,u_n(x))^t$ $A^t(x)\beta = a_2(x)(v_1(x),\cdots,v_n(x))^t$ where $a_i(x)$, $u_j(x)$ and $v_j(x)$ are polynomials such that $u_1,\cdots,u_n$ $(v_1,\cdots,v_n)$ do not have a common zero. For x fixed in the neighborhood of ξ $A(x)$ is a rank 1 matrix. So $A(x) = (a(x)u_i(x)v_j(x))_1^n$ where $a(x)$ is a rational function. It is left to show that $a(x)$ is polynomial. Suppose that $a(x) = b(x)/c(x)$ and $b(\eta) \neq 0$, $c(\eta) = 0$. As $A(x) \in M_n(\mathbf{C}[x])$ $u_i(\eta)\ v_j(\eta) = 0$ for $i,j = 1,\cdots,n$. Since $u_i(x)$, $i = 1,\cdots,n$ do not have a common zero there exists $1 \leq i \leq n$ such that $u_i(\eta) \neq 0$. Hence $v_j(\eta) = 0$, $j = 1,\cdots,n$, which contradicts our assumption that $v_1(x),\cdots,v_n(x)$ do not have a common zero. Hence $a(x)$ is a polynomial. Suppose that $A(x)$ is symmetric. Then we can choose $\beta = \alpha$ so $v(x) = u(x)$. ∎

Let $U(x) \in GL_n(\mathbf{C}[x])$, such that $U(0) = I$. Define

$$A(x) = U(x)A_0U^{-1}(x). \tag{8.13}$$

Then the eigenvalues of $A(x)$ are constant (do not depend on x). The converse of this statement is true if $A(x)$ has n distinct eigenvalues.

<u>Theorem</u> 8.14. Let $A(x) \in M_n(\mathbf{C}[x])$. Assume that $A(x)$ has constant pairwise distinct eigenvalues. Then (8.13) holds with $U(x) \in GL_n(\mathbf{C}[x])$ and $U(0) = I$. If in addition $A(x)$ is symmetric then $U(x) \in O_n(\mathbf{C}[x])$.

Proof. By considering the matrices $TA(x)T^{-1}$, $T \in GL_n$ we may assume that $A_0 = D = \mathrm{diag}\{d_1,\cdots,d_n\}$. As $A(x)$ has a constant spectrum each d_i is an eigenvalue of $A(x)$. Let

$$P_i(x) = \prod_{1\leq j\leq n, j\neq i} [d_jI - A(x)]/(d_j-d_i). \tag{8.15}$$

So $P_i(x)$ is rank one matrix with $\mathrm{tr}(P_i(x)) = 1$. Theorem 8.11 yields

$$P_i(x) = (u_j^i(x)v_k^i(x))_{j,k=1}^n, \quad \sum_{j=1}^n u_j^i(x)v_j(x) = 1.$$

Clearly $u^i(x) = (u_1^i(x),\cdots,u_n^i(x))^t$, $v^i(x) = (v_1^i(x),\cdots,v_n^i(x))^t$ are the eigenvectors of $A(x)$ and $A^t(x)$ respectively corresponding to the eigenvalue d_i. As $d_i \neq d_j$ for $i \neq j$ we have

$$\sum_{k=1}^{n} u_k^i(x)v_k^j(x) = 0 \quad \text{for } i \neq j.$$

Define

$$U(x) = (u^1(x),\cdots,u^n(x)), \quad V(x) = (v^1(x),\cdots,v^n(x)).$$

Then $V^t(x)U(x) = I$. So $U(x),V(x) \in GL_n(\mathbf{C}[x])$. As $u^i(x)$ is the eigenvector of $A(x)$ corresponding to d_i we have the equality

$$A(X)U(x) = U(x)D.$$

Also the assumption that $A_0 = D$ means that we can choose $u^i(x)$ to satisfy $u^i(0) = (\delta_{i1},\cdots,\delta_{in})^t$, $i = 1,\cdots,n$. This proves the theorem for a general $A(x)$. Suppose that $A(x) \in S_n(\mathbf{C}[x])$. Then we can choose $T \in O_n$. According to Theorem 8.11 $V = U$. So $U(x) \in O_n(\mathbf{C}[x])$. ∎

Theorem 8.14 does not apply if $A(x)$ has multiple eigenvalues. Indeed let

$$A(x) = \begin{pmatrix} x & ix \\ -ix & -x \end{pmatrix}.$$

Then $d_1 = d_2 = 0$ are the eigenvalues of $A(x)$. Also $A_0 = 0$. So (8.13) does not hold. Theorem 8.14 together with the equality (8.8) yield Theorem 7.15 for $n = 2$. Of course the case $n = 1,2$ can be proved easily in the direct way. However for the cases $n = 3,4$ (in particular $n = 4$) we need also the following characterization of $A(x) \in M_n(\mathbf{C}[x])$ with a constant spectrum. As usual let $[A,B]$ denote the commutant $AB-BA$.

Theorem 8.16. Let $F(x)$ be of the form (8.5). Consider the equation (8.2) with the initial condition $U(0) = I$. Let $A(x)$ be given by (8.13). Then $A(x)$ is of the form (8.1) if and only if $F(x)$ satisfies the following nonlinear equation of order m

$$[F^{(m)},A_0]+\cdots+[F^1,\underbrace{[F,\cdots,[F,A_0)\cdots]}_{m-1} + \underbrace{[F,[F,\cdots,[F,A_0]\cdots]}_{m+1} = 0. \tag{8.17}$$

In particular

$$\underbrace{[F_\ell,[F_\ell,\cdots,[F_\ell,A_0]\cdots]}_{m+1} = 0 \tag{8.18}$$

<u>Proof</u> The matrix $V(x) = U^{-1}(x)$ satisfies the equation (8.3). Thus if $A(x)$ is given by (8.13) we have

$$\frac{dA}{dx} = UFA_0V - UA_0FV = U[F,A_0]V$$

$$\frac{d^2A}{dx^2} = U\{[F^1,A_0] + [F,[F,A_0]]\}V \tag{8.19}$$

$$\frac{d^sA}{dx^s} = U\{[F^{(s-1)},A_0] + \cdots + \underbrace{[F,[F,\cdots,[F,A_0]\cdots]}_{s}\}V.$$

Thus $\frac{d^{m+1}}{dx^{m+1}} A(x) \equiv 0$ if and only if (8.17) holds. Assume that $F(x)$ is of the form (8.5). Then the coefficient of $x^{\ell(m+1)}$ in the left hand side of (8.17) is equal to $\underbrace{[F_\ell,[F_\ell,\cdots,[F_\ell,A_0]\cdots]}_{m+1} = 0$ This proves (8.18). ■

Assume that F_ℓ satisfies (8.18). By letting $F(x) \equiv F_\ell$ in (8.2) we obtain $F(x)$ which satisfies (8.17). Thus we proved

<u>Theorem</u> 8.20. Let F_ℓ satisfies (8.18). Then

$$A(x) = e^{F_\ell x}A_0e^{-F_\ell x} = A_0 + \sum_{k=1}^{m} \underbrace{[F_\ell,\cdots,[F_\ell,A_0]\cdots]}_{k}\frac{x^k}{k!}. \tag{8.21}$$

In particular $\underbrace{[F_\ell,\cdots,[F_\ell,A_0]\cdots]}_{m}$ is a nilpotent matrix.

The last assertion of the theorem follows from the result below.

Lemma 8.22. Let $A(x)$ be a nonconstant matrix of the form (8.1). If $A(x)$ has a constant spectrum then A_m is a nilpotent matrix.

Proof. As $tr(A^k(x))$ is constant we must have $tr(A_m^k) = 0$, $k = 1,2,\cdots$. So A_m is nilpotent. ■

Theorem 8.23. Let $A \in S_n$ have pairwise distinct eigenvalues. Then there exists $0 \neq B \in S_n$ such that $A+xB$ has a constant spectrum if and only if there exist $F \in A_n$ such that

$$[F,[F,A]] = 0, \quad [F,A] \neq 0. \tag{8.24}$$

In particular $A + x[F,A]$ has a constant spectrum, whence $[F,A]$ is nilpotent.

Proof. Let $A(x) = A+xB$. According to Theorem 8.14 the equality (8.13) holds for some $U(x) \in O_n(\mathbf{C}[x])$. As $B \neq 0$ $U(x)$ is not a constant matrix. Let $F(x)$ be given by (8.2). Then $0 \neq F(x) \in A_n(\mathbf{C}[x])$. So $F(x)$ is of the form (8.5) and F_ℓ is a nonzero skew symmetric matrix. In view of Theorem 8.16 $F = F_\ell$ satisfies the equality $[F,[F,A]] = 0$. Consider the matrix (8.21).

The equalities (8.19) yield

$$A(x) = A + x[F,A]. \tag{8.25}$$

According to Theorem 8.20 $A(x)$ has a constant spectrum. Lemma 8.22 implies that $[F,A]$ is nilpotent. It is left to show that $[F,A] \neq 0$. Suppose that $[F,A] = 0$. That is F and A commute. Since A has n distinct eigenvalues $F = p(A)$ for some polynomial $p(\lambda)$. So F is symmetric! Thus $F = 0$ which is a contradiction. ■

Thus to prove Theorem 7.15 for $n = 3,4$ we have to show that the only skew symmetric solution to (8.24) is $F = 0$ when A is a symmetric matrix with pairwise distinct eigenvalues. To prove Theorem 7.15 for $n \geq 5$ it is enough to find a nontrivial solution to (8.24) with the above restriction on F and A.

9. The equation $[F,[F,A]] = 0$.

We first recall some known facts about complex skew symmetric matrices. See for example Gantmacher [1959, Ch. 11].

First, if A and B are similar skew symmetric matrices then A and B are othogonally similar. Second, if zero is an eigenvalue of a skew symmetric matrix F then in the system of elementary divisors of F all those of even degree corresponding to the eigenvalue zero are repeated an even number of times. Thus if F is a nonzero 3×3 skew symmetric nilpotent matrix then F is orthogonally similar to the matrix H given in (8.9). If F is a nonzero 4×4 skew symmetric nilpotent matrix then either F is orthogonally similar to E given by (8.9) or F satisfies

$$F^2 = 0, \quad \ker(F) = \text{Range}(F). \tag{9.1}$$

Theorem 9.2. Let $0 \neq F \in A_n$ and $A \in S_n$. Assume that F and $[F,A]$ are nilpotent and compute. Then for $n=3,4$ A has at least one multiple eigenvalue.

Proof. We break our proof into three cases.

(i) $n = 3$. Then we may assume that $F = H$ and H is of the form (8.9). Since H and $[H,A]$ commute they have a common eigenvector u

$$Hu = 0 = [H,A]u = HAu.$$

Since the eigensubspace of H is spanned by u we must have $Au = \lambda u$. Also $u^t u = 0$. If λ was a simple eigenvalue (i.e. a simple zero of $|xI-A| = 0$) then the symmetricity of A yields that $u^t u \neq 0$. So λ is a multiple eigenvalue.

(ii) Let $n = 4$ and assume that (9.1) holds. Then the assumption that $[F,[F,A]] = 0$ implies $FAF = 0$. So $\ker(F) \supseteq A\,\text{Range}(F) = A\ker(F)$. If all the eigenvalues of A are simple then $\mathbf{C}^4$ splits

$$\mathbf{C} = \ker(F) \oplus W$$

i.e. W is an orthogonal complement of $\ker(F)$. So $FW \subseteq W$. Note that $\dim W = \dim \ker(F) = 2$. Also the eigenvectors of A, $\{x_i\}_1^4$ form an orthogonal basis of $\mathbb{C}^4$. In this basis F is represented by $F_1 \oplus F_2$ where each F_i is 2×2 skew symmetric matrix. Since F is nilpotent $F_i = 0$ so $F = 0$. This contradicts the assumption that $F \neq 0$. Whence A has a multiple eiganvalue.

(iii) Let $n = 4$ and assume

$$F = \operatorname{diag}(H,0) \qquad A = \begin{pmatrix} B & \beta \\ \beta^t & a \end{pmatrix},$$

where H is given by (8.9) and $B \in S_3$.

A straightforward calculation shows that

$$[F,[F,A]] = \begin{pmatrix} [H,[H,B]] & H^2\beta \\ \beta^t H^2 & 0 \end{pmatrix}.$$

So

$$[H,[H,B]] = 0, \quad \beta^t H^2 .$$

According to Theorem 8.20 $[H,B]$ is a nilpotent matrix. Then by the part (i) of this proof $Bu = \lambda u$ where u is the eigenvector of H. Since the range of H^2 is spanned by u the equality $\beta^t H^2 = 0$ implies that $\beta^t u = 0$. So

$$Av = \lambda v, \quad v^t = (u^t,0), \quad v^t v = u^t u = 0.$$

Thus λ is a multiple eigenvalue of A. ■

Theorem 8.23 and the above theorem imply Theorem 7.15 for $n = 3,4$. So we get

Theorem 9.3. Let $p(\lambda,z)$ be a nondegenerate polynomial of the form (5.1) with $r = 1$. Let $2 \leq n \leq 4$. Then there exists a symmetric pencil $A+xB$ whose characteristic polynomial is $p(\lambda,x)$. The number of such non-orthogonally similar pairs is at most

$$N(n) = \prod_{i=2}^{n} i!/2^{n-1} \tag{9.4}$$

and for almost all $p(\lambda,x)$ the number of distinct orbits of symmetric pairs is exactly $N(n)$.

Proof. By making a change of variables $x = x_1 + y$ and considering $Q(A+xB)Q^t$, $Q \in O_n$ we may assume that $p(\lambda,0)$ is of the form (7.12) and $A = D = \mathrm{diag}\{d_1,\cdots,d_n\}$, $B = E$. For $n \leq 4$ we showed that the system (7.14) has the unique solution $E = 0$. By Theorem 3.20 the map (7.13) is $\prod_{i=2}^{n} i!$ covering of $C^{(n+1)n/2}$. Thus there exist "generically" $\prod_{i=2}^{n} i!$ distinct symmetric E such that

$$|\lambda I-(D+xE)| = \phi(\lambda,x). \tag{9.5}$$

Obviously for most $p(\lambda,x)$ of the form (5.1) satisfying $p(\lambda,0) = |\lambda I-D|$ all the corresponding E have non-zero entries. Also $D+xE$ and $D+xPEP$, $P \in DO_n$ have the same characteristic polynomial. As E has nonzero entries then $P_1EP_1 = P_2EP_2$, $P_i \in DO_n$ if and only if $P_1 = P_2$. So each 2^{n-1} distinct E satisfying (9.5) belong to the same orbit. Thus, generically, there are $N(n)$ distinct orbits and we established the theorem.

For $n=2$ we obtain that if $p(\lambda,x)$ is non-degenerate then there exists only one orbit of symmetric pairs corresponding to $p(\lambda,x)$. This result was already obtained in Section 7. For $n=3,4$ we see that in general to a given $p(\lambda,x)$ of the form (5.1) with $r = 1$ correspond more than one orbit. In fact, it is possible to generalize Theorem 9.3 as follows.

Theorem 9.6. Let $D = \mathrm{diag}\{d_1,\cdots,d_n\}$ be a diagonal matrix with pairwise distinct diagonal entries. Let $p(\lambda,x)$ be a polynomial of the form (5.1) with $r = 1$ such that $p(\lambda,0) = |\lambda I-D|$. Assume that $2 \leq n \leq 4$. Let $B \in M_n$ be given. Then there exists $E \in S_n$ such that

$$|\lambda I - x(E+B)| = p(\lambda,x). \tag{9.7}$$

The number of such E never exceeds

$$M(n) = \prod_{i=2}^{n} i! \tag{9.8}$$

and for most of such $p(\lambda,x)$ the number of such E is exactly $M(n)$.

Proof. Consider the map

$$\varphi : S_n \to C^{(n+1)n/2} = P^{n,1}$$

given by

$$\varphi(E) = (P_{10}(D,E+B),p_{11}(D,E+B),\cdots,P_{nn}(D,E+B)).$$

So

$$\varphi_\pi(E) = (P_{10}(D,E),\cdots,P_{nn}(D,E)).$$

Thus $\varphi_\pi(E) = 0$ implies $E = 0$ and the result follows by Theorem 3.20. ∎

We now prove the second part of Theorem 7.15. By repeating the arguments of part (iii) in the proof of Theorem 9.2 we get

Lemma 9.9. Let $F \in A_n$ be orthogonally similar to E given by (8.9). Assume that $A \in S_n$ and $[F,[F,A]] = 0$. Then A has a multiple eigenvalue.

Thus in case that $n \geq 5$ we shall take E to be orthogonally similar to F of the form

$$F = \mathrm{diag}(H,0),\quad H = \begin{pmatrix} 0 & P \\ -P & 0 \end{pmatrix},\quad P = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}. \tag{9.10}$$

More precisely using orthogonal similarity we may assume that $A = A_1 \oplus A_2$ where $A_1 \in S_5$ and $F \in S$ of the form (9.10). Thus to prove Theorem

7.15 for $n \geq 5$ it is enough to consider $n = 5$. Choose

$$A = \begin{pmatrix} B & \beta \\ \beta^t & a \end{pmatrix}$$

where

$$B = \mathrm{diag}\{d_1 I, d_2 I\}$$

and I is 2×2 identity matrix. So

$$[F,[F,A] = \begin{bmatrix} [H,[H,B]] & H^2\beta \\ \beta H^2 & 0 \end{bmatrix} = \begin{bmatrix} [H,[H,B]] & 0 \\ 0 & 0 \end{bmatrix} .$$

Clearly

$$[H,[H,B]] = 0.$$

Choose $\beta^t = (0,b,0,c)$. Then $|\lambda I-A| = (\lambda-d_1)(\lambda-d_2)|\lambda I-C|$

$$C = \begin{vmatrix} d_1 & 0 & b \\ 0 & d_2 & c \\ b & c & a \end{vmatrix}$$

So

$$|\lambda I-C| = \lambda^3 - t_1\lambda^2 + t_2\lambda - t_3$$

$$t_1 = d_1 + d_2 + a$$

$$t_2 = d_1d_2 + (d_1+d_2)a - (b^2+c^2)$$

$$t_3 = d_1d_2a - d_2b^2 - d_1c^2.$$

Since $d_1 \neq d_2$, t_1, t_2 and t_3 determine uniquely a, b^2 and c^2. Thus we can choose a,b,c such that

$$|\lambda I-C| = (\lambda-d_3)(\lambda-d_4)(\lambda-d_5),$$

where $d_i \neq d_j$ for $i \neq j$. This proves the second part of Theorem 7.5.

10. Symmetric Polynomial Matrices.

Let $A(x), B(x) \in S_n(\mathbb{C}[x])$. That is

$$(10.1) \qquad A(x) = \sum_{j=0}^{m} A_j x^j, \quad B(x) = \sum_{j=0}^{m} B_j x^j, \quad A_j, B_j \in S_n, \; j = 0,\cdots,m.$$

The matrix $A(x)$ and $B(x)$ are called orthogonally similar if

$$(10.2) \qquad B(x) = U(x)A(x)U^t(x),$$

and $U(x) \in O_n(\mathbb{C}[x])$. If $U(x)$ can be chosen to be constant, i.e. $U(x) \equiv U \in O_n$ then $A(x)$ and $B(x)$ are said to be strictly orthogonally similar. That is the matrices $\{A_j\}_0^m$ and $\{B_j\}_0^m$ are simultaneously orthogonally similar

$$(10.3) \qquad B_j = UA_jU^t, \quad j = 0,\cdots,m.$$

Clearly for $n \geq 3$ there are symmetric matrices $A(x)$ and $B(x)$ which are orthogonally similar but not strictly orthogonally similar. Indeed, choose $A(x) \equiv A_0$ and $U(x) \in O_n(\mathbb{C}[x])$ such that $B(x)$ given by (10.2) is a non-constant matrix. So $B_m \neq 0$ and $A_m = 0$ for some $m \geq 1$. Obviously A_0 and $B(x)$ are not strictly orthogonally similar. Consider a characteristic polynomial of $A(x)$. It is easy to see that this polynomial is of the form

$$(10.4) \qquad \begin{aligned} |\lambda I - A(x)| &= \lambda^n + \sum_{i=1}^{n} p_i(x)\lambda^{n-i}, \\ p_i(x) &= \sum_{j=0}^{im} p_{ij}(A_0,\cdots,A_m)x^j, \qquad i = 1,2,\cdots,n. \end{aligned}$$

In particular

$$(10.5) \qquad |\lambda I - A_m| = \lambda^n + \sum_{i=1}^{m} p_{i(mi)}(A_0,\cdots,A_m)\lambda^{n-i}.$$

If $A(x)$ and $B(x)$ are orthogonally similar (or even similar over $GL_n(\mathbb{C}[x])$) then $A(x)$ and $B(x)$ have the same characteristic polynomial $p(\lambda,x)$. We now give a simple condition on $p(\lambda,x)$ which ensure strict orthogonal similarity of $A(x)$ and $B(x)$ provided that $A(x)$ and $B(x)$ are orthogonally similar.

Theorem 10.6. Let $A(x), B(x) \in S_n(\mathbf{C}[x])$ be of the form (10.1). Assume that (10.2) holds for $U(x) \in O_n(\mathbf{C}[x])$. Let $p(\lambda,x)$ of the form (10.4) be the characteristic polynomial of $A(x)$. Assume that A_m has pairwise distinct eigenvalues. Then $U(x)$ is a constant matrix. That is $A(x)$ and $B(x)$ are strictly orthogonally similar.

To prove the theorem we need the following lemmas.

Lemma 10.7. Let $A,B \in S_n$ have the same pairwise distinct eigenvalues. Then there exist 2^n distinct orthogonal matrices U such that $B = UAU^t$.

Proof. Since A and B are orthogonally similar to the same diagonal matrix D it is enough to consider the case $A=B=D$. But then $D = UDU^t$ if and only if $U = \mathrm{diag}\{\pm 1,\cdots,\pm 1\}$ and the lemma is proved.

Let Ω be a domain in $\mathbf{C}^k$. Denote by $H(\Omega)$ the set of analytic functions in Ω. Then $S_n(H(\Omega))$ and $O_n(H(\Omega))$ will denote the set of symmetric and orthogonal matrices $A(x)$ and $U(x)$ for $x \in \Omega$ such that the entries of $A(x)$ and $U(x)$ are analytic functions in Ω.

Lemma 10.8. Let Ω be a simply connected domain in $\mathbf{C}^k$. Assume that $A(x)$, $B(x) \in S_n(H(\Omega))$. Suppose that for each $x \in \Omega$, $A(x)$ and $B(x)$ have the same pairwise distinct eigenvalues. Let

$$B(x_0) = U_0 A(x_0) U_0^t, \quad U_0 \in O_n, \tag{10.9}$$

for some $x_0 \in \Omega$. Then there exists a unique $U(x) \in O_n(H(\Omega))$ satisfying (10.2) such that $U(x_0) = U_0$.

Proof. First we note that any $U(x) \in O_n(H(\Omega))$ satisfying (10.2) is a solution of linear and quadratic equations

$$B(x)U(x) - U(x)A(x) = 0, \quad U(x)U^t(x) = I.$$

As at each point x we have exactly 2^n distinct solutions. The implicit function theorem implies that for any given $x_1 \in \Omega$ there exists $r_1 = r(x_1) > 0$ such that the above system has 2^n distinct analytic solutions $U(x)$ in the disc $|x-x_1| < r_1$. So $U(x)$ can be continued analytically on any

continuous curve $\Gamma \subset \Omega$. Thus, in general, we have a multivalued (with at most 2^n branches) analytic orthogonal matrix valued function $U(x)$ satisfying (10.2) with $U(x_0) = U_0$ at least for some branch of $U(x)$. The assumption that ω is simply connected implies that $U(x)$ is univalued, i.e. $U(x) \in O_n(H(\Omega))$.

The assumption that Ω is simply connected is crucial to the proof of Lemma 10.8. Indeed let

$$A(x) = B(x) = \begin{pmatrix} 1+x & i \\ i & -1-x \end{pmatrix}.$$

Then in the domain

$$\Omega = \{x,\ 0 < |x| < 2\}$$

$A(x)$ has two distinct eigenvalues.

Clearly

$$A(-1) = U_0 A(-1) U_0^t, \quad U_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then the solution $U(x)$ satisfying (10.2) with $B(x) = A(x)$ and the condition $U(-1) = U_0$ is of the form

$$U(x) = A(x)/\sqrt{x(2+x)}, \qquad \sqrt{-1} = i .$$

Now $U(x)$ is two valued in Ω since $\sqrt{x(2+x)}$ is two valued in Ω. On the other hand if we choose $U_0 = I$ then $U(x) = I$ and this solution is clearly single valued in Ω. This shows that the choice of U_0 is important.

<u>Proof of Theorem</u> 10.6. Let $|x| > r$. Then by dividing (10.2) by x^m we get

$$C(x^{-1}) = B_m + \sum_{i=1}^{m} B_{m-i} x^{-i} = U(x)(A_m + \sum_{i=1}^{m} A_{m-i} x^{-i}) U^t(x).$$

Put $y = x^{-1}$. For $|y| < \epsilon$ the eigenvalues of $C(y)$ would be close to the eigenvalues of A_m and therefore will be pairwise distinct. Thus we can apply Lemma 10.8 for $y^m A(y^{-1})$ and $y^m B(y^{-1})$. Hence $U(x)$ is analytic in the neighborhood of $x = \infty$. Now the Liouville's theorem implies that $U(x)$ must be constant. ■

Remark 10.10. The results of Theorem 10.6 apply if we shall assume that $U(x) \in O_n(H(\mathbf{C}))$, i.e. the entries of $U(x)$ are entire functions.

Let $A(x)$ be of the form (10.1). Suppose that $A(x)$ is non-degenerate. That is $|\lambda I-A(x)|$ is a non-degenerate polynomial. Hence $D(x)$ given by (4.17) is a polynomial of degree $\leq n(n-1)m$. Assume furthermore that A_m has pairwise distinct eigenvalues. Then $D(x)$ is a polynomial of degree $n(n-1)m$. Let $\zeta_1,\cdots,\zeta_s$ be all intersection points of $|\lambda I-A(x)|$. So $s \leq n(n-1)m$. Let $B(x)$ be also of the form (10.1) and suppose that

$$|\lambda I-A(x)| = |\lambda I-B(x)|. \tag{10.11}$$

Pick $x_0 \neq \zeta_j$, $j = 1,\cdots,s$. Let U_0 satisfy (10.9). Then, according to the proof of Lemma 10.8 there exists an analytic multivalued orthogonal matrix function $U(x,U_0)$ on the Riemann sphere (i.e. $\mathbf{C} \cup \{\infty\}$) punctured at ζ_j, $j = 1,\cdots,s$, satisfying (10.2) such that

$$U(x_0,U_0) = U_0$$

for at least one branch of $U(x,U_0)$. If it happens that $U(x,U_0)$ is single valued and the points ζ_j, $j = 1,\cdots,s$, are removable singularities then the Liouville's theorem implies that $U(x_0,U_0) = U_0$ and $A(x)$ and $B(x)$ are strictly orthogonally similar. Note, as we pointed out before, sometime this can happen only for special choices of U_0 out of 2^n possible choices. In what follows we give a simple condition on U_0 which will ensure the needed properties of $U(x,U_0)$.

Lemma 10.12. Let $A(x),B(x)$ be of the form (10.1). Assume that $A(x)$ and $B(x)$ have the same characteristic polynomial and suppose that A_m has pairwise distinct eigenvalues. (So $|\lambda I-A(x)|$ is non-degenerate). Let $\zeta_1,\cdots,\zeta_s$ be the intersection points of $|\lambda I-A(x)|$. Choose $x_0 \neq \zeta_j$, $j= 1,\cdots,s$ such that the real line $(1-t)x_0 + t\zeta_j$, t real contains only one intersection point ζ_j. Let $U_j(x)$ be the analytic solution of the equation

(10.2) along the open segment $(1-t)x_0 + t\zeta_j$, $0 \leq t < 1$ with the initial condition $U_j(x_0) = U_0$ for a fixed choice of U_0 satisfying (10.9). Assume that $U_j(x)$ can be continued analytically to a disc $|x-\zeta_j| < r$ for some $r > 0$ and $j = 1,\cdots,s$. Then $U_j(x) = U_0$, $j = 1,\cdots,s$, hence $A(x)$ and $B(x)$ are strictly orthogonally similar.

Proof. Let Ω be the complex plane cut along the rays $(1-t)x_0 + t\zeta_j$, $1 \leq t$, $j = 1,\cdots,s$. So Ω is simply connected, $\zeta_j \notin \Omega$, for $j = 1,\cdots,s$. Thus, there exists a unique $U(x) \in O_n(H(\Omega))$ satisfying (10.2) and the condition $U(x_0) = U_0$. Let

$$\hat{\Omega}_1 = \Omega \quad \{(1-t)x_0 + t\zeta_1,\ t > 1\}.$$

Note that $\zeta_1 \notin \hat{\Omega}_1$. We can continue $U(x)$ analytically along any closed curve Γ in $\hat{\Omega}_1$. We claim that $U(x)$ is a single valued in $\hat{\Omega}_1$. Indeed if Γ is homotopic to a point in $\hat{\Omega}_1$ then of course $U(x)$ is single valued along Γ. It is left to examine the case where Γ is homotopic to circling k times around the point ζ_1. Since $U_1(x)$ can be continued analytically in the neighborhood of ζ_1 we see that the $U(x)$ remains single valued on Γ in this case too. Finally, the assumption that $U_1(x)$ can be continued analytically in the neighborhood of ζ, implies that $U(x)$ is single valued and analytic in $\Omega_1 = \hat{\Omega}_1 \cup \{\zeta_1\}$. Continuing in the same manner we deduce that $U(x)$ can be extended analytically to the whole complex plane. As A_m has pairwise distinct eigenvalues we get that $q(\lambda)$ given by (10.5) has simple roots. Now Remark 10.10 implies that $U(x) \equiv U_0$. ■

Let $A(x)$, $B(x)$ be of the form (10.1). Suppose that $A(x)$ and $B(x)$ have the same characteristic polynomial which is non-degenerate. In order to be able to apply the above lemma we have to insure that $A(x)$ and $B(x)$ are (i) analytically similar in the neighborhood of each intersection point ζ. That is

$$(10.13) \qquad B(x) = U(x)A(x)U(x)^{-1}, \quad U(x) \in GL_n(H(D)).$$

Here D is some disc

$$(10.14) \qquad D(\zeta,r) = \{x, |x-\zeta| < r\}$$

and $GL_n(H(\Omega))$ is the set of invertible matrices $U(x)$ such that the entries of $U(x)$ and $U^{-1}(x)$ are analytic functions in Ω.

(ii) $U(x)$ can be chosen to be an orthogonal matrix.

The question of local (analytic) similarity was studied by us in Friedland [1980]. The first step is to bring $A(x)$ to a block diagonal form that

$$(10.15) \qquad V^{-1}(x)A(x)V(x) = \sum_{j=1}^{t} \oplus A_j(x), \quad V(x) \in GL_n(H(D(\zeta,r))),$$

such that $A_j(\zeta)$ has one eigenvalue λ_j, $j = 1,\cdots,m$ and $\lambda_j \neq \lambda_k$ for $j \neq k$. Here r is some positive number. Also

$$(10.16) \qquad W(x)^{-1}B(x)W(x) = \sum_{j=1}^{t} \oplus B_j(x), \quad W(x) \in GL_n(H(D(\zeta,e))),$$

and $A_j(x)$ and $B_j(x)$ have the same characteristic polynomial in $D(\zeta,r)$ for $j = 1,\cdots,t$. So $A(x)$ and $B(x)$ are locally analytically similar if and only if $A_j(x)$ and $B_j(x)$ are locally similar for $j = 1,\cdots,t$. Clearly if $A_j(x)$ and $B_j(x)$ are 1×1 matrices then $A_j(x) = B_j(x)$ since $A_j(x)$ and $B_j(x)$ have the same characteristic polynomial. In that case $A_j(x)$ and $B_j(x)$ are locally similar.

In case that $A_j(x)$ and $B_j(x)$ are not 1×1 matrices it may well happen that $A_j(x)$ and $B_j(x)$ have the same characteristic polynomial but $A_j(x)$ and $B_j(x)$ are not locally similar. A simple criterion due to Wasow [1963] gives an additional condition on $A_j(\zeta)$ and $B_j(\zeta)$ which ensures the local similarity of $A_j(x)$ and $B_j(x)$.

<u>Theorem</u> 10.17. Let $A(x), B(x) \in M_n(H(D(\zeta,r)))$. Assume that $A(x)$ and $B(x)$ have the same characteristic polynomial. If the minimal polynomial of $A(\zeta)$ and $B(\zeta)$ is equal to its characteristic polynomial then $A(x)$ and $B(x)$ are locally similar.

Let $A(x) \in M_2(H(D))$ for some disc (10.13). The interesting case from our point of view is when $A(\zeta)$ has a double eigenvalue μ_0. In the neighborhood of ζ the eigenvalues of $A(x)$ behave as an algebriac function of x, i.e. they must have a Puiseaux expansion (4.5) with $m = 2$. For a certain Puiseaux expansion of $\lambda(x)$ it is possible to tell when the minimal polynomial of $A(\zeta)$ is $(\lambda-\mu_0)^2$.

Lemma 10.18. Let $A(x) \in M_2(H(D))$, $D = D(\zeta,r)$. Assume that the eigenvalues of $A(x)$ have the Puiseaux expansion (4.5) with $m = 2$ in the neighborhood of ζ. If $\mu_1 \neq 0$ then the minimal polynomial of $A(\zeta)$ is $(\lambda-\mu_0)^2$.

Proof. Assume to the contrary that $A(\zeta) = \lambda_0 I$. So

$$(10.19) \qquad A(x) = \lambda_0 I + (x-\zeta)B(x), \quad B(x) \in M_2(H(D)).$$

Since the eigenvalues of $B(x)$ have the Puiseaux expansion we deduce that $\mu_1 = 0$ contrary to our assumptions. ∎

Let $A(x)$ be of the form (10.1) and suppose that A_m has pairwise distinct eigenvalues. So $|\lambda I-A(x)|$ is non-degenerated and the discriminent $D(x)$ given by (4.17) is a polynomial of degree $n(n-1)m$. Let

$$(10.20) \qquad \delta(A_0,\cdots,A_m) \equiv \delta(|\lambda I-A(x)|)$$

be the discriminant of $D(x)$ given by (5.9) $(r=m)$.

Suppose that

$$(10.21) \qquad \delta(A_0,\cdots,A_m) \neq 0.$$

Then at each intersection point ζ, $A(\zeta)$ has exactly one double root μ_0. Moreover the Puiseaux expansion of those two eigenvalues satisfy the assumptions of Lemma 10.18 (Theorem 4.22). So if $B(x)$ is of the form (10.1) and $|\lambda I-A(x)| = |\lambda I-B(x)|$ then $A(x)$ and $B(x)$ are locally similar in the neighborhood of any finite or infinite point ζ if the condition (10.20) holds. Our next step is to show that the similarity marix in (10.13) can be chosen to be orthogonal. This is implied by the following two lemmas.

<u>Lemma</u> 10.22. Let $A(x) \in S_n(H(D(\zeta,r')))$. Suppose that $A(\zeta)$ has m distinct eigenvalues $\lambda_1,\cdots,\lambda_m$, where n_j is the multiplicity of λ_j. Then there exists $V(x) \in O_n(H(D(\zeta,r)))$ satisfying (10.15) where $A_j(\zeta)$ has one eigenvalue λ_j for some $r > 0$.

<u>Proof.</u> By considering the matrix $QA(\zeta)Q^t$, $Q \in O_n$ we may assume that

$$A(\zeta) = \sum_{j=1}^{m} \oplus A_j(\zeta),$$

where each $A_j(\zeta)$ has one eigenvalue λ_j.

Choose a positive ρ such that

$$D(\lambda_i,\rho) \cap D(\lambda_j,\rho) = \phi \quad \text{for} \quad i \neq j.$$

Let

$$E_j(x) = \frac{1}{2\pi i} \int_{|\lambda-\lambda_j|=\rho} (\lambda I - A(x))^{-1} d\lambda$$

$$E_j^2(x) = E_j(x) = E_j^t(x), \quad E_j(x)E_k(x) = 0$$

$$\text{for} \quad k \neq j, \quad \text{rank } E_j(x) = n_j, \quad j = 1,\cdots,m$$

for $|x-\zeta| < r''$ for some positive r''.

Let

$$\ell_j = n_0+\cdots+n_{j-1}, \quad j = 1,\cdots,m+1, \quad n_0 = 0, \quad u_\alpha^t = (\delta_{\alpha 1},\cdots,\delta_{\alpha n})^t.$$

Consider

$$u_\alpha(x) = E_j(x)u_\alpha, \quad \alpha = \ell_j+1,\cdots,\ell_{j+1}.$$

Note that $u_\alpha(\zeta) = u_\alpha$, $\alpha = 1,\cdots,n$. Apply the Gram-Schmidt process to $u_{\ell_j+1}(x),\cdots,u_{\ell_{j+1}}(x)$ to obtain analytic vectors $v_{\ell_j+1}(x),\cdots,v_{\ell_{j+1}}(x)$ in the neighborhood of ζ such that

$$v_\alpha^t(x)v_\beta(x) = \delta_{\alpha\beta}, \quad \alpha,\beta = \ell_{j+1},\cdots,\ell_{j+1}, \quad j = 1,\cdots,m.$$

Since $E_j(x)E_k(x) = 0$ for $j \neq k$ we get $v_\alpha^t(x)\, v_\beta(x) = \delta_{\alpha\beta}$, $1 \leq \alpha,\beta \leq n$.

Then

$$V(x) = (v_1(x),\cdots,v_n(x)) \in O_n(H(D(\zeta,r)))$$

and the equality (10.15) holds. ∎

Lemma 10.23. Let $A(x) \in S_2(H(D(\zeta,r')))$. Then there exists $U(x) \in O_2(H(D(\zeta,r)))$ for some $r > 0$ such that $U(x)A(x)U^t(x)$ has one of the following forms

$$\text{(i)} \qquad \begin{pmatrix} a(x) & 0 \\ 0 & b(x) \end{pmatrix},$$

$$\text{(ii)} \qquad \begin{pmatrix} 1+a(x) & b(x) \\ b(x) & -1+a(x) \end{pmatrix}, \; b(\zeta) = i$$

$$\text{(iii)} \qquad \left(\sum_{j=0}^{k} \alpha_j (x-\zeta)^j\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (x-\zeta)^{k+1}\begin{pmatrix} 1+a(x) & b(x) \\ b(x) & -1+a(x) \end{pmatrix},$$

$$b(\zeta) = i, \quad k \geq 0.$$

In all the cases the functions $a(x), b(x)$ and the values $\alpha_1,\cdots,\alpha_k$ are determined by the characteristic polynomial of $A(x)$. If the eigenvalues of $A(x)$ are not identical then the value of k in (iii) is bounded from above.

Proof. Suppose first that $A(\zeta)$ has two distinct eigenvalues. Then $A(x)$ is orthogonally similar to the matrix of the form (i) by virtue of Lemma 10.22. Assume now that $A(\zeta)$ has a double eigenvalue. By considering the matrix

$$A(x) - \tfrac{1}{2}\operatorname{tr}(A(x))I$$

we may assume that $A(x)$ of the form

$$A(x) = \begin{pmatrix} \alpha(x) & \beta(x) \\ \beta(x) & -\alpha(x) \end{pmatrix}.$$

Suppose that the minimal polynomial of $A(\zeta)$ is λ^2. Using the transformation $QA(x)A^t$, $Q \in O_2$ we may assume $\alpha(\zeta) = 1$, $\beta(\zeta) = i$. Then choose in the form (ii)

$$b = \beta\sqrt{1 + (\alpha^2-1)\beta^{-2}}$$

and let

$$U(x) = \begin{pmatrix} u(x) & -v(x) \\ v(x) & u(x) \end{pmatrix}, \quad v = \frac{(\alpha-1)}{\beta+b}u, \quad u = 1/\sqrt{1+(\alpha-1)^2(\beta-b)^{-2}}.$$

Then $U(x)A(x)U^t(x)$ will be of the form (ii) with $a(x) \equiv 0$. Assume now that the minimal polynomial of $A(\zeta)$ is λ. Then $A(\zeta) = 0$ and

$$A(x) = (x-\zeta)B(x), \quad B(x) \in S_2(H(D(\zeta,r))).$$

If the minimal polynomial of $B(\zeta)$ is λ^2 then $B(x)$ is of the form (ii) and we achieved the form (iii). If $B(\zeta) = 0$ then we continue in the same manner. If we stop the process at the step k we shall achieve the form (iii). If the above process never stops then we deduce that $A(x)$ is of the form (i) and $a(x) \equiv b(x)$. It is easy to see that $a(x), b(x), \alpha_1, \cdots, \alpha_k$ are determined by $tr(A(x))$ and $|A(x)|$. Suppose we have the form (iii). Then the Puiseaux expansion of both eigenvalues of $A(x)$ are of the form

$$\lambda_h(x) = \sum_{j=0}^{k} \alpha_j (x-\zeta)^j + |x-\zeta|^{k+1} O(1), \quad h=1,2.$$

Thus if $\lambda_1(x) \not\equiv \lambda_2(x)$, k has an upper bound. ∎

Let $A(x)$ be of the form (10.1). Suppose that $\delta(A_0, \cdots, A_m) \neq 0$. Then $D(x)$ given by (4.17) is a polynomial of degree $n(n-1)m$ having $n(n-1)m$ simple zeros $\zeta_1, \cdots, \zeta_s$, $s = n(n-1)m$. Thus $|\lambda I - A(x)|$ has $n(n-1)m$ intersection points $\zeta_1, \cdots, \zeta_s$. At each intersection point ζ_k there exists $V_k(x) \in O_n(H(D(\zeta_k,r)))$ such that (10.15) holds. Here $t = n-1$, $A_1(x) \in S_2(H(D(\zeta_k,r)))$, $A_1(x)$ is of the form (ii)

$$(10.24) \qquad A_1(x) = C_k(x) = \begin{pmatrix} 1 + a_k(x) & b_k(x) \\ b_k(x) & -1+a_k(x) \end{pmatrix}, \quad b_k(\zeta_k) = i$$

in view of Theorem 4.22, Lemmas 10.18 and 10.23. Other $A_j(x)$ are the analytic roots of $|\lambda I - A(x)|$ in the neighborhood of ζ_0. Assume that $B(x)$ is of the form (10.1) and suppose that $|\lambda I - A(x)| = |\lambda I - B(x)|$. Then at the neighborhood of each intersection point ζ_j there exists $W_k(x) \in O_n(H(D(\zeta_k,r)))$ such that (10.16) holds and

$$(10.25)\qquad \sum_{j=1}^{n-1} \oplus A_j(x) = \sum_{j=1}^{n-1} \oplus B_j(x).$$

Choose $x_0 \in \mathbb{C}$, $x_0 \neq \zeta_j$, $j=1,\cdots,n(n-1)m$, such that the line $(1-t)x_0 + t\zeta_j$, $-\infty < t < \infty$, contains only one intersection point ζ_j. By considering the matrices $Q_1A(x)Q_1^t$, $Q_2B(x)Q_2^t$, $Q_j \in O_2$, $j = 1,2$ we assume

$$(10.26)\qquad A(x_0) = B(x_0) = D = \mathrm{diag}\{d_1,\cdots,d_n\}.$$

Let Ω_k be a simply connected domain Ω_k containing the segment $(1-t)x_0 + t\zeta_k$, $0 \leq t < 1$ and such that $\zeta_j \in \Omega_k$, $j = 1,\cdots,s$. Choose $T \in DO_n$. Let $U_k(x,T) \in O_n(H(\Omega_k))$ be the unique solution of (10.2) satisfying $U_k(x_0,T) = T$. Thus for $x \in \Omega_k \cap D(\zeta_k,r) = \hat{\Omega}_k$ we have

$$B(x) = W_k(x)\left(\sum_{j=1}^{n-1} \oplus A_j(x)\right)W_k^t(x) = U_k(x,T)A(x)U_k^t(x,t) =$$

$$U_k(x,T)V_k(x)\left(\sum_{j=1}^{n-1} \oplus A_j(x)\right)V_k^t(x)U_k^t.$$

As $A_j(\zeta)$ and $A_\ell(\zeta)$ do not have a common eigenvalue for $j \neq \ell$ we easily deduce

$$(10.27)\qquad W_k^tU_kV_k = \sum_{j=1}^{n-1} \oplus R_j \in O_n(H(\hat{\Omega}_k)).$$

In particular $|R_j(x)|^2 = 1$. So for $j \geq 2$ each $R_j(x)$ is a constant number of modulus 1.

For $R_1(x)$ we have

$$(10.28)\qquad \begin{aligned} &R_1(x) = Q_k(x,B,T) \in O_2(H(\hat{\Omega}_k)) \\ &Q_k(x,B,T)C_k(x)Q_k^t(x,B,T) = C_k(x) \end{aligned}$$

where $C_k(x)$ is given by (10.24).

Let

$$(10.29)\qquad \omega_k(B,T) = |Q_k(x,B,T)|,\quad k = 1,\cdots,n(n-1)m.$$

So $\omega_k(B,T) = \pm 1$.

<u>Lemma</u> 10.30. Let the above assumption hold. Suppose that $\omega_k(B,T) = 1$ then the matrix $U_k(x)$ can be continued analytically to the neighborhood of ζ_k.

Proof. Consider the matrix $C_k(x)$, $x \in \hat{\Omega}_k$. So $C_k(x)$ has two distinct eigenvalues. THus there are four orthogonal matrices commuting with $C_k(x)$. Namely $\pm I$, $\pm Q$, where $|Q| = -1$. Therefore if $\omega_k(B,T) = 1$ then $Q_k(x,B,T)$ is a constant matrix which equals to $\pm I$.

Hence $\sum_{j=1}^{n-1} \oplus R_j(x)$ is constant in $\hat{\Omega}_k$ and therefore can be analytically extended to $D(\zeta_k,r)$. Finally (10.27) shows that

$$U_k(x,T) = W_k(x)\left(\sum_{j=1}^{n-1} \oplus R_j(x)\right)V_k^t(x)$$

so $U_k(x,T)$ has analytic extension to $D(\zeta_k,r)$. ■

Theorem 10.31. Let $A(x), B(x), \breve{B}(x) \in S_n(\mathbf{C}[x])$ be of the form (10.1). Assume that

$$|\lambda I - A(x)| = |\lambda I - B(x)| = |\lambda I - \breve{B}(x)|.$$

Suppose that $\delta(A_0,\cdots,A_m) \neq 0$. Let $\zeta_1,\cdots,\zeta_s$, $s = n(n-1)m$ be the intersection points of $|\lambda I - A(x)|$ and suppose that the line $(1-t)x_0 + t\zeta_j$, $-\infty < t < \infty$, only one intersection point ζ_j for $j = 1,\cdots,s$.

Assume that $A(x_0) = B(x_0) = \breve{B}(x_0)$ is a diagonal matrix. Let $T, \breve{T} \in DO_n$ and suppose that

$$\omega_k(B,T) = \omega_k(\breve{B},\breve{T}), \quad j = 1,\cdots,n(n-1)m. \tag{10.32}$$

Then

$$\breve{B}(x) = \breve{T}\, TB(x)T\breve{T}. \tag{10.33}$$

Proof. In the arguments preceding Lemma 10.30 replace B,U,W,R by $\breve{B},\breve{U},\breve{W},\breve{R}$ respectively. So we get

$$\breve{U}_k = \breve{W}_k\left(\sum \oplus \breve{R}_j(x)\right)V_k^t(x).$$

Thus

$$\breve{U}_k U_k^t = \breve{W}_k\left(\sum_{j=1}^{n-1} \oplus \breve{R}_j(x)R_j^t(x)\right)W_k^t$$

and

$$\breve{R}_1(x)R_1^t(x)C_k(x)R_j\breve{R}_j^t(x) = C_k(x).$$

As

$$|\tilde{R}_j(x)R_j^t(x)| = \omega_k(\tilde{B},\tilde{T})\omega_k(B,T) = 1$$

we deduce that $\tilde{R}_1(x)R_1^t(x) = \pm I$ so $\tilde{R}_1(x)R_1^t(x)$ has an analytic continuation in the neighbnorhood of ζ_k. Hence the orthogonal matrix $A(x)$ satisfying the equation

$$\tilde{B}(x) = Q(x)B(x)Q^t(x), \quad Q(x_0) = \tilde{T}T$$

fulfills the assumptions of Lemma 10.12. So $Q(x) = \tilde{T}T$ and (10.33) holds.

<u>Theorem</u> 10.34. Let $p(\lambda,x)$ be a polynomial of the form (10.4). Assume that $\delta(p) \neq 0$. Then there are at most $\nu = 2^{(n-1)(mn-1)}$ polynomial symmetric matrices $A_1(x),\cdots,A_\nu(x)$ of the form (10.1) such that $|\lambda I-A_j(x)| = p(\lambda,x), \quad j = 1,\cdots,\nu$ and

$$A_j(x) \neq QA_i(x)Q^t$$

for $i \neq j$ and any $Q \in O_n$.

<u>Proof.</u> Assume to the contrary that the matrices $A_1(x),\cdots,A_\nu(x)$ and $A(x)$ are of the form (10.1), their characteristic polynomial is $p(\lambda,x)$ and any of the matrices are not strictly orthogonally similar. Choose x_0 as in Theorem 10.31. We assume that

$$A(x_0) = A_j(x_0) = D, \quad j = 1,\cdots,\nu$$

where D is a diagonal matrix. We first note that if $T_1,T_2 \in OD_n$ and $T_1 \neq T_2$ then

$$\omega_k(A_j,T_1) \neq \omega_k(A_j,T_2)$$

for some k. Otherwise Theorem 10.31 imply that

$$A_j = T_1T_2A_jT_1T_2.$$

As $T_1T_2 \neq \pm I$ it easily follows that there exists a permutation matrix such that

$$PA_jP^t = A_{j1} \oplus A_{j2}.$$

Therefore

$$p(\lambda,x) = |\lambda I - A_{j1} \oplus A_{j2}| = |\lambda I_1 - A_{j1}||\lambda I_2 - A_{j2}|$$

and this is impossible since $p(\lambda,x)$ is irreducible.

Also for any $T_1, T_2 \in OD_n$

$$\omega_k(A_i, T_1) \neq \omega_k(A_j T_2), \qquad i \neq j$$

for some k. Otherwise $A_i(x)$ and $A_j(x)$ are strictly orthogonally equivalent which contradicts our assumptions. Let

$$\omega(A_i, T) = (\omega_1(A_i, T), \cdots, \omega_s(A_i, T)) \qquad s = n(n-1)m.$$

Then the set $\{\omega(A_i, T)\}$, $i=1, \cdots, \nu$, $T \in OD_n$ contains $\nu\, 2^{n-1} = 2^{(n-1)nm}$ distinct vectors. As $\omega_k(A_i, T) = \pm 1$, $k = 1, \cdots, n(n-1)m$ we deduce that there exists i such that $\omega_k(A_i, T) = 1$, $k = 1, \cdots, n(n-1)m$. Clearly $\omega_k(A, I) = 1$, $k=1, \cdots, n(n-1)m$. Thus Theorem 10.31 implies that $A(x)$ and $A_i(x)$ are strictly orthogonally equivalent which contradicts our assumptions. The proof of the theorem is completed.

11. Conclusions, remarks and open problems.

Let $\mathcal{S}_{n,m+1}$ be the space of $m+1$ symmetric tuples $(A_0,\cdots,A_m)$

(11.1) $$\mathcal{S}_{n,m+1} = \underbrace{S_n \times \cdots \times S_n}_{m+1} .$$

Thus $\mathcal{S}_{n,m+1}$ can be identified with $\mathbb{C}^{(n+1)n(m+1)/2}$. The complex orthogonal group O_n is acting naturally on $\mathcal{S}_{n,m+1}$

(11.2) $$U[(A_0,\cdots,A_m)] = U(A_0,\cdots,A_m)U^t, \quad U \in O_n .$$

To each $m+1$ symmetric tuple $(A_0,\cdots,A_m)$ we correspond a polynomial symmetric matrix $A(x) \in S_n(C[x])$ by means of the formula (10.1). Consider the characteristic polynomial $|\lambda I-A(x)|$ given by (10.4). Clearly $p_{ij}(A_0,\cdots,A_m)$ is a polynomial on $\mathcal{S}_{n,m+1}$ which is invariant under the action of O_n. So $p_{ij}(A_0,\cdots,A_m) \in [\mathcal{S}_{n,m+1}]^0$. The coefficients of $|\lambda I-A(x)|$ induce the map

(11.3) $$p:\mathcal{S}_{n,m+1} \to P^{n,m+1},\ p(A_0,\cdots,A_m) = (p_{ij}(A_0,\cdots,A_m)),$$

$$i = 1,\cdots,n, \quad j = 0,\cdots,mi,$$

where $P^{n,r}$ is the coefficient space of the polynomials of the form (5.1). Denote by $\delta(p(A_0,\cdots,A_m))$ the discriminant of $|\lambda I-A(x)|$ given by (5.9) $(r=m)$. We claim that $\delta \not\equiv 0$ on $\mathcal{S}_{n,m+1}$. The proof is identical to the proof of Theorem 6.5 except that instead of the pencil $A+xB$ we have to consider the polynomial matrix $A+x^mB$. Next we claim that the transcendence degree of $[\mathcal{S}_{n,m+1}]^0$ is $mn(n+1)/2 + n$. Indeed, assume that A_0 has pairwise distinct eigenvalues. So there exist $m+1$ tuple $(D,E_1,\cdots,E_m)$ orthogonally similar to $(A_0,\cdots,A_m)$ such that D is a diagonal matrix. The matrices $E_1,\cdots,E_m$ are fixed up to the action of the discrete group DO_n. Thus $(D,E_1,\cdots,E_m)$ parametrize most of the orbits in $[\mathcal{S}_{n,m+1}]^0$. So the transcendence degree of $[\mathcal{S}_{n,m+1}]^0$ is at most $mn(n+1)/2+n$. An obvious modification of Theorem 7.4 yields

Theorem 11.4. The polynomials

(11.5) $\quad \mathrm{tr}(A_0^i),\ i = 1,\cdots,n,\ \mathrm{tr}(A_0^i A_k A_0^j A_k),\quad 0 \le i < j \le n-1,\quad k = 1,\cdots,m$

form a transcendence basis in $[\mathcal{S}_{n,m+1}]^0$. Moreover, if A_0 has n distinct eigenvalues then the values of these polynomials determine at most $2^{m(n^2-n+2)/2}$ distinct orbits.

Theorem 10.34 implies

Theorem 11.6. The polynomials $p_{ij}(A_0,\cdots,A_m)$, $i = 1,\cdots,n$, $j = 0,\cdots,mi$, given by (10.4) form a transcendence basis in $[\mathcal{S}_{n,m+1}]^0$. More precisely, if $\delta(p(A_0,\cdots,A_m)) \ne 0$ then the values of these polynomials determine at most $2^{(n-1)(mn-1)}$ distinct orbits.

Proof. As the transcendence degree of $[\mathcal{S}_{n,m+1}]^0$ is $mn(n+1)/2+n$ it is enough to show that $p_{ij}(A_0,\cdots,A_m)$ $i = 1,\cdots,n$, $j = 0,1,\cdots,im$ are algebraically independent. That is the map p given in (11.3) is proper. Suppose that p is not proper. Then for "most" of $\pi \in p(\mathcal{S}_{n,m+1})$, $p^{-1}(\pi)$ is a variety of dimension

$$\frac{n(n-1)}{2} + 1 = (m+1)n(n+1)/2 - [mn(n+1)/2+n]+1$$

at least. Choose $(A_0,\cdots,A_m) \in \mathcal{S}_{n,m+1}$ such that A_0 has pairwise distinct eigenvalues and $\delta(p(A_0,\cdots,A_m)) \ne 0$. According to Theorem 10.34 there are at most $2^{(n-1)(mn-1)}$ distinct orbits satisfing $p(B_0,\cdots,B_m) = \pi = p(A_0,\cdots,A_m)$. As the eigenvalues of B_0 are pairwise distinct, there are only 2^n orthogonal matrices which commute with B_0 So the orbit of $(B_0,\cdots,B_m)$ under the action of O_n is of dimension $n(n-1)/2$ - the dimension of the connected component of O_n. Therefore $p^{-1}(\pi)$ has dimension $n(n-1)/2$ for such $(A_0,\cdots,A_m)$. Clearly, the set of $m+1$ symmetric tuples $(A_0,\cdots,A_m)$ such that $\delta(p(A_0,\cdots,A_m)) \ne 0$ and A_0 has a pairwise distinct eigenvalue is an open (algebraical) set in $\mathcal{S}_{n,m+1}$. So, for "most" of $\pi \in p(\mathcal{S}_{n,m+1})$, $p^{-1}(\pi)$ is a variety of dimension $n(n-1)/2$. The above contradiction proves that p is a proper map. The proof of the theorem is completed. ∎

We conjecture

Conjecture 11.7. The map (11.3) is onto map.

The results of Sections 7-9 confirm the conjecture for $n \leq 4$ and $m \leq 2$. Theorem 11.6 yields

$$(11.8) \qquad \deg p \leq 2^{(n-1)(mn-1)}.$$

Problem 11.9. Find the degree of the map p given by (11.3).

Next we observe that the results of Theorem 10.34 apply to a larger class of matrices.

Theorem 11.10. Let $(A_0,\cdots,A_m) \in \mathcal{S}_{n,m+1}$ and consider $A(x) \in S(C[x])$ given by (10.1). Assume

(i) for each $\zeta \in C$, each eigenvalue of $A(\zeta)$ is either simple or double,

(ii) each eigenvalue of A_m is either simple or double.

Then the values of the polynomials $p_{ij}(A_0,\cdots,A_m.$, $i = 1,\cdots,n$, $j = 0,\cdots,mi$, (i.e. the characteristic polynomial $|\lambda I-A(x)|$) determine at most $2^{(n-1)(mn-1)}$ distinct orbits.

Proof. The assumptions of the theorem imply that $p(\lambda,x) = |\lambda I-A(x)|$ is non-degenerate. Moreover for each intersection point ζ, $A(x)$ is orthogonally similar to $\sum_{j=1}^{t} \oplus A_j(x)$, i.e. V in (10.15) belongs to $O_n(H(D(\zeta,r)))$, where each $A_j(x)$ is either 2×2 or 1×1 symmetric matrix. Suppose that $A_j \in \mathcal{S}_2(H(D(\zeta,r)))$. Since $p(\lambda,x)$ is not degenerate $A_j(x)$ has two distinct eigenvalues $\lambda_\alpha(x) \not\equiv \lambda_\beta(x)$. According to Lemma 10.23 $A_j(x)$ can be analytically similar to a finite number of matrices of the form (i)-(iii) and this number can be bounded by using the Puiseaux series of the eigenvalues of $A(x)$ at $x = \zeta$. Therefore we have a finite number of classes $\mathcal{A}_1,\cdots,\mathcal{A}_t$ of $A(x)$ of the form (10.1) such that

(i) all of them have the same characteristic polynomial $p(\lambda,x)$

AD-A114 536 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
SIMULTANEOUS SIMILARITY OF MATRICES.(U)
MAR 82 S FRIEDLAND DAAG29-80-C-0041
UNCLASSIFIED MRC-TSR-2345 NL

2 OF 2
AD A114536

END
DATE
FILMED
6-82
DTIC

(ii) if $A(x), B(x) \in \mathcal{A}_j$, then $A(x)$ and $B(x)$ are orthogonally similar in the neighborhood of any ζ (finite or infinite). Now, the arguments in the previous sections show that in each $\mathcal{A}_j$ there is only a finite number $A_1(x),\cdots,A_{k_j}(x)$ such that $A_\alpha(x)$ and $A_\beta(x)$ are not strictly orthogonally similar for $1 \leq \alpha < \beta < k_j$. So $p_{ij}(A_0,\cdots,A_m)$, $i = 1,\cdots,n$, $j = 0,\cdots,mi$, determine a finite number of distinct orbits. The inequality (11.8) combined with Theorem 3.17 implies the number of distinct orbits determined by $p_{ij}(A_0,\cdots, A_m)$, $i = 1,\cdots,n$, $j = 0,\cdots,mi$ is at most

$2^{(n-1)(mn-1)}$. ∎

In many physical applications $A_0,\cdots,A_m$ are symmetric and real. That is we restrict ourselves to $\mathcal{S}_{n,m+1}(R)$. In that case the real orthogonal group $O_n(R)$ is acting naturally on $\mathcal{S}_{n,m+1}(R)$. In particular each orbit $\mathrm{orb}(A_0,\cdots,A_m)$ is compact. We pose the following problem.

Problem 11.11. Let $(A_0,\cdots,A_m) \in \mathcal{S}_{n,m+1}(R)$. Does the characteristic polynomial $|\lambda I - \sum_{i=0}^{m} A_i x^i|$ determine always a finite number of orbits? What is the value of this number?

As in Section 5 let $\mathcal{M}_{rd}$ denote the variety of reducible polynomials in $P^{n,m+1}$. Put

$$(11.12) \qquad \mathcal{S}_{rd,n,m+1} = p^{-1}(\mathcal{M}_{rd}).$$

If p was onto map then Corollary (5.22) would yield

$$(11.13) \qquad \mathrm{codim}\mathcal{S}_{rd,n,m+1} \leq m(n-1).$$

Problem 11.14. Find $\mathrm{codim}\mathcal{S}_{rd,n,m+1}$ and $\mathrm{codim}\mathcal{S}_{rd,n,m+1}(R)$ in $\mathcal{S}_{n,m+1}(R)$. Here $\mathcal{S}_{rd,n,m+1}(R) = \mathcal{S}_{rd,n,m+1} \cap \mathcal{S}_{n,m+1}(R))$.

In Friedland-Simon [1981] we showed

$$(11.15) \qquad \mathrm{codim}\mathcal{S}_{rd,n,1}(R) \leq n-1$$

and we conjectured the equality sign in (11.15). We proved this conjecture

for $n=2,3$. We now point out briefly how to prove this conjecture for $n=4$

(11.16) $$\operatorname{codim} \mathcal{S}_{rd,4,1}(R) = 3$$

Assume that $A_0, A_1 \in \mathcal{S}_4(R)$ and let $p(\lambda,x) = |\lambda I-(A_0+xA_1)|$. Assume first that $p(\lambda,x)$ is not degenerate. Then Theorem 9.6 implies that there is only a finite number of $\mathrm{orb}(A_0,A_1)$ with the characteristic polynomial $p(\lambda,x)$. Recall that in this case $\mathrm{orb}(A_0,A_1)$ is a real variety of dimension $6 = 4 \cdot 3/2$ Assume next that $p(\lambda,x)$ is degenerate. It can be shown in Shapiro [1979] that for $n \leq 4$ the Kippenhahn conjecture [1951] is valid.

Conjecture 11.17. (Kippenhahn). Let A,B be $n \times n$ Hermitian matrices. Assume that $A+xB$ is a degenerate pencil. Then A and B have a common non-trivial subspace. That is there exists a unitary matrix U such that

$$U^*(A+xB)U = (A_1+xB_1) \oplus (A_2+xB_2).$$

In case that A,B are real symmetric U can be chosen to be a real orthogonal matrix.

Thus, it follows that $p(\lambda,x)$ in this case determines also a finite number of orbits $\mathrm{orb}(A_0,A_1)$ whose dimension is 6 at most. Since the codimension of $\mathcal{M}_{rd}(R)$ in $P^{n,1}(R)$ is 3 we obtain the equality (11.6). Also the above arguments show that Problem 11.11 has a positive answer for $m=1$ and $n \leq 4$.

REFERENCES

Avron, J. E. and B. Simon [1978], Analytic properties of band functions, Annals of Physics 110, 85-101.

Brenner, S. [1975], Quivers with commutativity conditions and some phenomenology of forms, Lect. Notes Math. 488, 25-53.

Engel, M. and H. Schneider [1973], Cyclic and diagonal products on a matrix, Lin. Alg. and Appl. 7, 301-335.

Fogarty, J. [1969], Invariant Theory, Benjamin, New York.

Friedland, S. [1977], Inverse eigenvalue problems, Lin. Alg. Appl. 17, 15-51.

Friedland, S. [1980], Analytic similarity of matrices, Lectures on Applied Math,, vol. 18, 43-85, Amer. Math. Soc, Providence, R.I.

Friedland, S. [1981], Classification of orbits under the action of an algebraic group on varieties, submitted.

Friedland, S. and B. Simon [1981], The codimension of degenerae pencils, Lin. Alg. and Appl., to appear.

Gantmacher, F. [1959], The Theory of Matrices, vol. I and II, Chelsea, New York.

Gelfand, I.M. [1970], The cohomology of infinite dimensional Lie algebras, some questions fo integral geometry, Actes, Congrés Intern. Math. vol. 1, 95-111.

Gelfand, I.M. and V. A. Ponomarev [1969], Remarks on the classification of a pair of commuting linear transformations in a finite dimensional vector space, Funktsional'nyi Analiz ievo Prilozheniya 3, 81-82 [English translation: Functional Analysis and Its Applications, 3 (1969), 325-326].

Kippenhahn, R. [1951], Uber den wertevorrat einer matrix, Math. Nachr. 6, 193-228.

Lutz, D. A. [1967], On systems of linear differential equations having regular singular solutions, J. Differential Eq. 3, 311-322.

Marcus, M. and H. Minc [1964], A Survey of Matrix Theory and Matrix Inequalities, Prindle, Weber and Schmidt, Boston.

Moiseyev, N. and S. Friedland, The association of resonance states with incomplete spectrum of finite complex-scaled Hamiltonian matrices, Physical Reviews A22, 618-624.

Nathanson, M. [1980], Classification problems in K-categories, Fundamenta Mathematicae 15, 187-197.

Noether, E. and B. L. van der Waerden [1928], Die alternative bei nichtlinearen gleichungen, Nachrichten der Gesellschaft der Wissenschaften zu Göttingen Math.-Phys. Klasse, 77-87.

Processi, C. [1976], The invariant theory of $n \times n$ matrices, Advances in Math. 19, 306-381.

Shafarevich, I. R. [1974], Basic Algebraic Geometry, Springer-Verlag, New York,

Shapiro, H. [1979], Unitary block diagonalization and the characteristic polynomials of a pencil generated by Hermitian matrices, Ph.D. Dissertation, Cal.Tech.

van der Waerden, B. L. [1950], Modern Algebra, vol. I and II, Unger, New York

Whitney, H. [1972], Complex Analytic Varieties, Addison-Wesley, Reading, Massachusetts

Wasow, W. [1963], On holomorphically similar matrices, J. Math. Anal. and Appl. 4, 202-206.

SF/db

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2345	2. GOVT ACCESSION NO. AD-A114 536	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SIMULTANEOUS SIMILARITY OF MATRICES		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Shmuel Friedland		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 1 - Applied Analysis
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE March 1982
		13. NUMBER OF PAGES 96
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Simultaneous similarity, invariant functions, symmetric matrices

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we solve completely and explicitly the long standing problem of classifying pairs of $n\times n$ complex matrices (A,B) under the simultaneous similarity (TAT^{-1},TBT^{-1}). Roughly speaking, the classification breaks into a finite number of steps. In each step we consider an open algebraic set $M^0_{n,2,r,\rho} < M_n \times M_n$ (M_n - the set of $n\times n$ complex valued

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract (continued)

matrices). Here r and ρ are two positive integers. Then we construct a finite number of rational functions $\phi_1,\cdots,\phi_s$ in the entries of A and B whose values are constant on all pairs similar in $\mathcal{M}_{n,2,r,\rho}$ to (A,B). The values of the functions $\phi_i(A,B)$, $i = 1,\cdots,s$, determine a finite number (at most $\kappa(n,2,r)$) of similarity classes in $\mathcal{M}_{n,2,r,\rho}$. Let S_n be the subspace of complex symmetric matrices in M_n. For $(A,B) \in S_n \times S_n$ we consider the similarity class (TAT^t, TBT^t) where T ranges over all complex orthogonal matrices. Then the characteristic polynomial $|\lambda I - (A+xB)|$ determines a finite number of similarity classes for almost all pairs $(A,B) \in S_n \times S_n$.